



HAL
open science

Les politiques éducatives et le pouvoir du chiffre : les évaluateurs nationaux des résultats du système éducatif à l'épreuve de PISA
La Direction de l'évaluation, de la prospective et de la performance du ministère de l'Éducation nationale et l'émergence d'une évaluation internationale des résultats des élèves, le Programme International pour le suivi des acquis des élèves (1995 – 2017)

Anaïs Le Bouffant

► **To cite this version:**

Anaïs Le Bouffant. Les politiques éducatives et le pouvoir du chiffre : les évaluateurs nationaux des résultats du système éducatif à l'épreuve de PISA La Direction de l'évaluation, de la prospective et de la performance du ministère de l'Éducation nationale et l'émergence d'une évaluation internationale des résultats des élèves, le Programme International pour le suivi des acquis des élèves (1995 – 2017). Science politique. Université Paris sciences et lettres, 2022. Français. NNT : 2022UPSLD009 . tel-04633706

HAL Id: tel-04633706

<https://theses.hal.science/tel-04633706v1>

Submitted on 3 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE DE DOCTORAT

DE L'UNIVERSITÉ PSL

Préparée à l'Université Paris Dauphine

**Les politiques éducatives et le pouvoir du chiffre : les
évaluateurs nationaux des résultats du système
éducatif à l'épreuve de PISA**

La Direction de l'évaluation, de la prospective et de la performance du ministère de l'Éducation nationale et l'émergence d'une évaluation internationale des résultats des élèves, le Programme International pour le suivi des acquis des élèves (1995 – 2017)

Soutenue par

Anaïs DUBREUCQ LE

BOUFFANT

Le 1^{er} février 2022

Ecole doctorale n°ED 543

Ecole doctorale SDOSE

Spécialité

Science politique

Composition du jury :

Dorota DAKOWSKA

Professeure de Science politique à Sciences Po Aix

Rapporteure

Marine DE LASSALLE

Professeure de sociologie à Sciences Po Strasbourg

Rapporteure

Emmanuel HENRY

Professeur de Sociologie à l'Université Paris Dauphine-PSL

Président du jury

Etienne PENISSAT

Chargé de recherche au CERAPS UMR 8026

Examineur

Eric AGRIKOLIANSKY

Professeur de Science politique à l'Université Paris Dauphine-PSL

Directeur de la thèse

Résumé

En 2000 est lancée l'évaluation PISA sous l'égide de l'OCDE, à laquelle participe la France. La DEPP, direction des évaluations statistiques interne au ministère de l'Éducation nationale, à l'institutionnalisation récente, est chargée d'en suivre la mise en œuvre pour la France et de garantir la validité des résultats pour aider au pilotage de la politique éducative nationale. Dans un contexte où l'évaluation des résultats de l'action éducative est encore peu instituée, l'irruption de PISA dans le panel des évaluations à disposition des décideurs publics remet en question le rôle de producteur de données pour l'action que la DEPP s'est tout juste vue confier. A la fois interlocutrice privilégiée de l'OCDE pour PISA mais sans réel pouvoir d'agir en profondeur sur les choix de méthode et d'objectifs faits par l'organisation internationale, en charge d'évaluations nationales encore peu utilisées pour piloter l'action publique, la DEPP va devoir repenser les caractéristiques de son expertise pour s'assurer que ses évaluations et ses analyses sont bien employées à cette fin, complétées par les données internationales de PISA et non l'inverse. Afin d'observer le repositionnement des experts de l'évaluation statistique en éducation au moment de la construction de PISA et jusqu'à la publication relative à son cycle 2016, cette thèse s'intéresse à la construction de l'identité d'expert en évaluation des résultats de l'action éducative de la DEPP et de l'OCDE avec PISA. Elle cherche ensuite à mettre en lumière les stratégies de relégitimation de ses travaux d'évaluation menées par les agents de la DEPP, entre critique méthodologique de PISA et intérêt philosophique et technique de l'objet. Enfin, elle observe la stratégie à l'épreuve de l'usage, en regardant la façon dont les résultats PISA et les évaluations de la DEPP sont employées dans le cadre du contrôle de l'action du gouvernement et lors des deux grandes réformes de la période, l'instauration du socle commun de compétences en 2005 et la refondation de l'école pour lutter contre les inégalités scolaires en 2013.

Mots-clés : Evaluation des politiques éducatives, organisation internationale, service statistique ministériel, expertise, PISA, OCDE, ministère de l'Éducation nationale, concurrence des acteurs, résultats de l'éducation, fabrique de la loi, pilotage de l'action publique, évaluation standardisée, entrepreneur de problème public.

Abstract

In 2000, PISA is launched under the auspices of the OECD, with the participation of France. The DEPP, the internal statistical evaluation directorate of the Ministry of education, which had recently been institutionalized, is put in charge of monitoring its implementation in France and ensuring the validity of the results to help steer the national education policy. The irruption of PISA as a tool for policy makers to evaluate education policies in France, where the evaluation of education policies is in the process of being institutionalised, calls into question the recently acquired role of the DEEP as data producer. The Directorate is the OECD's direct interlocutor, but has no real ability to influence the objectives and choices of method made by the international organisation. The Directorate is at the same time in charge of the national education evaluations and assessments which are still little used to steer public action. The DEPP will thus have to rethink its expertise to ensure that its evaluations and assessments, as well as its analyses of the results, are used for policy definition, and are complemented by PISA data, not the other way around. In order to study the repositioning of the DEEP and the OECD as experts in the evaluation of education policy, this thesis analyses the process of construction of such role using PISA, from the construction of international assessment until its 2016 cycle. It then analyses the strategies for legitimising DEEP's role as education policy evaluator, highlighting the paradox of their methodological criticism of PISA and their philosophical and technical interest for it. Finally, it looks at that strategy in action, observing how PISA results and DEPP evaluations and assessments are used for the control of governmental action and in the two major reforms of that period, the establishment of the common base of skills in 2005 and the 2013 "refounding of the school" aiming at tackling school inequities.

Keywords : education policies assessment, international organisation, ministerial service of statistics, expertise, PISA, OECD, ministry of education, competition between actors, education outcomes, policy making , public action steering, standardised evaluation, public problem entrepreneur.

Remerciements

Me voici arrivée au bout d'un processus de sept années. Sept années durant lesquelles cette thèse aura été un lien avec la joie des études, celle de « creuser » un sujet et de s'inscrire dans un temps long, trop long au goût de certains, mais finalement le juste temps pour moi.

Ces sept ans auront été le temps de quatre expériences professionnelles passionnantes dont ces travaux ont largement bénéficié, par les connaissances immenses qu'elles m'ont apporté et les personnes que j'ai pu y croiser, mais aussi par le recul qu'elles m'ont permis d'avoir sur mon objet de recherche et sur le processus de recherche en lui-même.

Ils auront été le temps de vivre dans deux pays, dans trois villes, autant de risques de ne plus retrouver le chemin de la thèse qui ont été heureusement évités grâce à la patience et les encouragements toujours compréhensifs de mon directeur de thèse Eric Agrikoliansky. Je n'aurais jamais pu parcourir ce chemin tortueux et aboutir ce travail sans son accompagnement et ses conseils riches et précieux, je l'en remercie infiniment. Nous aurons expérimenté les réunions de travail et les comités de thèse en visioconférence avant le covid, à 10 000 km de distance et avec des connections hasardeuses. Merci de cette souplesse et de cette confiance, je suis sûre que cela n'a pas été toujours évident

L'entrée en thèse fût difficile, sortant d'une dernière formation peu conventionnelle pour s'engager dans un travail de thèse en sciences politiques. Sur les conseils d'Olivier Nay, que je remercie pour son soutien à ma volonté de mener un doctorat en parallèle de mon activité professionnelle et pour son optimisme quant au succès de mes démarches, j'ai contacté Eric Agrikoliansky qui a cru en mon projet. Le laboratoire de recherche IRISSO, que j'ai à regret peu fréquenté, a suivi mes lentes avancées et mes demandes de réinscription un peu trop nombreuses avec l'exigence requise. Même si les rituels de réinscription ont parfois été complexes, j'imagine bien la difficulté à croire en la capacité de

terminer une thèse de quelqu'un que l'on n'a pour ainsi dire jamais vu et jamais écouté. Je remercie les responsables de l'Ecole Doctorale de m'avoir finalement accordé la confiance nécessaire pour finaliser mes travaux.

Mes recherches de doctorat ont parfois été solitaires. Je veux remercier Elise Tenret et Emmanuel Henry pour leurs lectures attentives, leurs critiques toujours constructives et les échanges agréables et dynamisants que nous avons pu avoir lors des différents comités de suivi de thèse qui ont été organisés ces dernières années. Ils ont été finalement mon lien avec le monde doctoral, et j'ai eu beaucoup de chance d'avoir pu construire mes travaux avec leur soutien.

Une thèse n'est rien sans le matériau qui a contribué à la réaliser. Je veux bien sûr remercier tous ceux qui m'ont accordé de leur temps, souvent avec un agenda très chargé, parfois à plusieurs reprises, pour que nous puissions nous entretenir sur les différents questionnements qui ont guidé mes recherches. Merci également à toutes celles et ceux avec qui j'ai pu avoir des moments de discussions informelles qui ont enrichi mes réflexions et mes hypothèses de travail, de la salle des conférences du Sénat aux couloirs de l'OCDE. Cette thèse porte en elle toutes les rencontres professionnelles qui ont jalonné le début de ma carrière, et qui sont pour certaines devenues de vraies amitiés me questionnant toujours avec bienveillance sur l'avancée de mes travaux.

Si finalement mes nombreux passages aux archives nationales auront été peu fructueux pour diverses raisons, ceux aux archives de l'OCDE m'auront en revanche fourni une quantité inattendue d'archives précieuses qui ont alimenté largement cette thèse. Je tiens à remercier le service des archives de l'OCDE pour leur accueil et l'aide qu'ils ont pu m'apporter pour trouver ces pépites.

Le parcours du combattant de la thèse, encore plus lorsqu'elle se prolonge, aurait été impossible à mener sans les encouragements et les petites attentions de mes amis, Sophie, Charlotte, Sarah, Sarah, Camille, et Cécile en particulier, de mes sœurs, de mes parents, de mes grands-parents. Merci pour vos petits mots, vos

pensées dans les périodes de creux et de travail, vos encouragements quand le temps et l'énergie venaient à manquer. Ils ont été toujours là au bon moment.

Enfin, la thèse faisant partie du quotidien pendant de longues années, je veux te dire Rodrigo toute mon admiration pour ta patience et tous mes remerciements pour avoir été présent à mes côtés, écouté mes plaintes, mes doutes, mes agacements. Tes talents de contradicteur, ton œil de relecteur, nos discussions sur PISA agrémentées parfois de débats sur les attendus universitaires français versus ceux de monde anglo-saxons ont fait le succès de cette entreprise.

Léon, quel bonheur d'avoir partagé cette dernière année et demi de thèse avec toi, de Paris à Lyon en passant par la Bretagne et Angers. Elle m'a sans doute pris du temps, mais les souvenirs des après-midi d'écriture en te regardant dormir en rendent la mémoire plus douce.

Sommaire

Introduction générale	11
Partie 1 : La construction de PISA, un paradigme à plusieurs égards dans l'histoire des évaluations des politiques éducatives françaises....	41
Chapitre 1 : L'évaluation en éducation : la naissance tardive d'un besoin de gestion publique en France	43
1) Une histoire longue de l'évaluation de l'action éducative.....	46
2) La statistique comme méthode au centre de la constitution du champ des évaluateurs d'Etat.....	75
Chapitre 2 : La création de PISA : « l'œil du Persan » ou l'externalisation de l'expertise des politiques éducatives	99
1) L'entrée de l'OCDE dans le théâtre de l'internationalisation de l'expertise en politique éducative	101
2) La genèse d'une évaluation qui se veut inédite	115
3) La nouveauté PISA : un instrument d'action publique inédit pour évaluer les résultats de l'éducation	140
Partie 2 : L'irruption de nouveaux agents « experts » de l'évaluation des politiques éducatives : un repositionnement en deux temps des évaluateurs nationaux.....	157
Chapitre 3 : Évaluateurs nationaux et internationaux : une même mission à partager ?	159
1) Les évaluateurs nationaux au cœur de nouveaux paradigmes de l'évaluation de l'action éducative.....	161
2) DEPP et OCDE: le « fonctionnaire savant » et le « fonctionnaire expert »	176
3) Une évaluation PISA dans les faits largement orientée par l'OCDE	204
Chapitre 4 : La cristallisation des enjeux d'influence à travers l'usage de l'outil.....	223
1) L'invisibilisation de l'expertise de la DEPP à l'origine de critiques de fond de PISA	230
2) Faire de PISA un faire-valoir des évaluations nationales de la DEPP	256

Partie 3: La stratégie à l'épreuve de l'usage dans la fabrique de la loi et le contrôle de l'action de l'Etat.	303
Chapitre 5 : Les parlementaires et les évaluations statistiques de l'action éducative : usage politique des données et expertise de ses producteurs	307
1) Quelles évaluations pour mettre à l'agenda les problèmes publics ?	314
2) Comparer: usage différencié des évaluations et internationalisation des défis de l'éducation nationale	331
3) L'enjeu de l'indépendance des évaluations et donc de leurs évaluateurs.....	349
Chapitre 6 : PISA et la réforme éducative : l'expertise internationale au cœur de la fabrique de la loi nationale	373
1) La construction de l'argument fondateur de la réforme : quels évaluateurs pour quelle expertise ?	380
2) L'usage de l'instrument dans le processus de construction de la mise en place de la volonté réformatrice	404
3) Le temps long de la réforme : l'expertise au service de la construction de l'évaluation des résultats de la réforme	435
Conclusion générale.....	453
Bibliographie	467
Ouvrages.....	467
Articles	474
Rapports et manuels.....	487
Lois, textes règlementaires, débats parlementaires, lettres de mission, communications institutionnelles	493
Colloques.....	498
Sites internet	498
Annexes.....	501
Annexe 1. Liste des entretiens réalisés	501
Annexe 2. Note anonyme rédigée en interne à la DEPP à propos de PISA	504

Introduction générale

Le 14 décembre 2020, le président de la République française Emmanuel Macron prend la parole depuis le Château de la Muette, siège parisien de l'Organisation de Coopération et de Développement Economique (OCDE), pour en célébrer le 60^e anniversaire. Devant le secrétaire général Angel Gurría, le président du gouvernement d'Espagne Pedro Sanchez, le président du Conseil européen Charles Michel, la directrice générale de l'UNESCO Audrey Azoulay, et un parterre d'ambassadeurs, il loue les grands succès historiques de l'OCDE et la méthode qu'elle emploie pour construire ses travaux jugés d'une qualité irréprochable car ayant fait les preuves de leur robustesse et de leur utilité politique:

« Cette convention de 1960, que nous célébrons, portait en elle tous les succès de l'OCDE que nous connaissons aujourd'hui. Ces succès reposent sur une méthode éprouvée, remarquable, autour de trois piliers : d'abord des données rigoureuses permettant à chaque pays de se comparer, ensuite des échanges de meilleures pratiques entre experts, et enfin des recommandations générales et adaptées à chaque État membre. Des grands succès ont été construits grâce à cette méthode. (...) Je pense par exemple à PISA dans le domaine de l'éducation qui nous aide à mesurer, à faire progresser nos niveaux scolaires. Je pense bien sûr aux travaux contre l'évasion fiscale, à la fin du secret bancaire, à la lutte contre la corruption. Je pense aussi au rapport de l'OCDE sur les financements du climat en 2015, qui a contribué au succès de la COP21 et de l'Accord de Paris. Je pense aux standards de l'OCDE sur la conduite responsable des entreprises. (...) Je crois dans la science, les faits, l'expertise indépendante plutôt que dans le complotisme et la manipulation de l'information ou le nouvel obscurantisme. Face à cela, je crois dans la coopération libre, choisie, intelligente plutôt que dans le repli nationaliste ou la brutalité. Ces deux croyances se trouvent pleinement incarnées par votre organisation et Monsieur le Secrétaire général, par ce que vous avez su conduire durant ces 14 années ici à Paris, dans le monde entier pour nous tous.»

L'évaluation PISA, acronyme du Programme international pour le suivi des acquis des élèves, est citée en tête des travaux phares de l'OCDE au moment du

bilan de ses 60 années d'existence. Cela reflète les motifs initiaux ayant présidé à la création d'une telle organisation à la sortie de la Seconde Guerre Mondiale : un projet politique international, nourri de la coopération économique entre les pays membres et fondé sur une expertise scientifique indépendante, pour « *rendre les vies meilleures* ¹ ». L'organisation internationale se voit reconnue un statut d'expert indépendant, censé fournir en toute transparence et objectivité des données et préconisations accompagnant le pilotage des politiques nationales, notamment éducatives, des pays membres.

Le président français est l'un des seuls chefs d'État membre à prendre la parole à l'occasion de cette célébration. Certes, cette allocution presque solitaire est liée en partie à des raisons relatives à la structure même de l'organisation et au contexte sanitaire. Le siège de l'OCDE est après tout à Paris, ce qui confère à la France et à son plus haut représentant politique une place particulière lors des moments de communication sur la vie de l'organisation. Cette proximité géographique est d'autant plus importante pour expliquer cette introduction française que le monde connaît depuis un an une pandémie liée au virus du covid-19. La crise sanitaire incitant en cette année 2020 à réduire drastiquement les déplacements et les regroupements, elle peut expliquer le petit nombre de chefs d'Etats et de gouvernements prenant la parole pour l'occasion.

Néanmoins, l'intervention du Président français constitue un indice fort de l'importance prise par l'OCDE dans l'espace des politiques publiques en France et en particulier de l'évaluation PISA, qui joue ces dernières années un rôle croissant dans le pilotage de l'action publique en éducation en France. L'évaluation des résultats de l'éducation dans les pays qui y participent est citée avant les projets menés par l'organisation internationale dans son champ de compétence historique, la coopération économique et financière. Alors que l'éducation ne fait pas *a priori* parti des missions constitutives de l'organisation, créée après-guerre pour déployer les fonds du plan Marshall et que l'évaluation

¹ Le slogan de l'OCDE est le suivant : « *de meilleures politiques pour des vies meilleures* », en anglais « *Better policies for better lives* ».

en elle-même n'a « que » 20 ans d'existence, PISA a acquis une popularité et une légitimité qui dépasse tout autre publication OCDE.

Dans le cas du pilotage de la politique éducative française plus précisément, qui sera le cadre géographique d'analyse de nos travaux tout au long de cette thèse, l'influence de PISA dans la conduite des politiques publiques nationale n'a cependant pas toujours été la même. A son démarrage à la fin des années 1990, l'évaluation est connue et reprise seulement par quelques « initiés ² » travaillant à l'évaluation de la politique éducative française. Les résultats PISA sont le plus souvent ignorés ou cibles de critiques méthodologiques fortes remettant en cause sa significativité lorsqu'appliquée au système éducatif français. Pourtant, la France, pays d'accueil du siège de l'OCDE, participe bien à chaque cycle de PISA, et cela depuis sa construction puis son lancement au tout début des années 2000. A première vue, le constat en 2010, au moment où naissent les premiers questionnements préparatoire à cette thèse, est frappant : comment la France peut-elle investir autant de moyens depuis presque 10 ans pour participer aux cycles PISA et « simplement » réfuter de toute part la recevabilité de ses résultats ? Dix années plus tard, PISA est au contraire convoqué régulièrement par les acteurs du champ des politiques éducatives, qu'ils en soient les concepteurs, les évaluateurs, les cibles ou encore les commentateurs. L'intervention d'Emmanuel Macron en est une illustration parlante.

Cette thèse interroge cette évolution des usages de PISA dans les discours des dirigeants politiques français, pour ce qu'elle signifie du rôle d'expertise alloué publiquement à l'évaluation de l'OCDE. En miroir, elle observe ce que cette arrivée d'un nouvel évaluateur-expert international, historiquement peu investi dans le champ des évaluations des résultats de l'éducation, implique pour les évaluateurs institutionnels nationaux dont c'est la mission centrale. Dans le cas français, ces évaluateurs sont membres de la DEPP, la Direction de l'évaluation,

² L'expression est employée par Nathalie Mons et Xavier Pons, « Pourquoi n'y a-t-il pas eu de « choc PISA » en France ? Sociologie de la réception d'une enquête internationale (2001-2008) », *Revue française de Pédagogie*, n°182, Voir aussi Jean-Claude Emin, « Les enquêtes internationales en France : de l'indifférence à la référence », *Administration et éducation*, n°3, 2012.

de la prospective et de la performance³ du ministère de l'Éducation nationale et par ailleurs interlocuteurs directs de l'OCDE pour la mise en place des cycles PISA auprès des élèves français et la validation des résultats produits par l'OCDE. Ce rôle de médiateur qui leur est alloué par le gouvernement au lancement du projet PISA à partir du milieu des années 1990 intervient au moment où leur rôle d'évaluateur statistique national s'est tout juste institutionnalisé, avec difficulté. Cette mission nouvelle est par ailleurs menée dans le cadre du lancement d'un outil d'évaluation d'un genre nouveau, qui n'a pas été développé pour correspondre aux caractéristiques spécifiques du système éducatif français, mais dont les conclusions sont en pratique devenues omniprésentes dans les discours des décideurs politiques au moment de contrôler ou de réorienter les politiques éducatives nationales.

Ce rôle de PISA dans la conduite de la politique éducative française révèle l'influence de l'expertise de ceux qui la produisent, et questionne le positionnement des évaluateurs nationaux à la fois en matière de capacité à se constituer entrepreneurs de problèmes publics, ce dont ils se réclament, mais aussi à légitimer leur mission d'évaluateur indépendant de l'action du gouvernement, crédible pour contribuer au pilotage de l'action publique et au contrôle du pouvoir exécutif. Entre adhésion philosophique et rejet méthodologique, cette évaluation des systèmes éducatifs, qui paraissait n'être qu'une parmi les autres, semble avoir eu un effet à la fois sur les interactions entre politiques éducatives et évaluation pour ce qui concerne les dispositifs, et sur les interactions entre évaluateurs et avec leurs rapports aux décideurs politiques pour ce qui concerne les acteurs.

³ Par souci de simplicité, nous nommerons dans cette thèse la DEPP par son dernier acronyme en date, quelle que soit le nom qu'elle portait à la période évoquée. Seul le chapitre 1 retrace l'histoire de la DEPP et reprend donc les différents noms qu'elle a pu porter. La DEPP a en effet porté plusieurs noms une fois institutionnalisée comme direction ministérielle. Elle s'est appelée à sa création DEP, direction de l'évaluation et de la prospective. En 1998, à l'arrivée au ministère de Claude Allègre, elle est renommée DPD, Direction de la programmation et du développement. Elle retrouvera sa dénomination de direction d'évaluation statistique en 2005, année où elle devient Direction de l'évaluation, de la prospective et de la performance (DEPP). Voir le chapitre 1 pour plus de détails, notamment sur le changement des missions allouées à la direction au moment des changements de dénomination.

Plusieurs questions se posent alors, auxquelles cette thèse veut répondre : d'abord, comment l'arrivée de nouveaux experts-évaluateurs internationaux dans le champ des politiques éducatives modifie-t-elle le rôle des experts-évaluateurs nationaux, producteurs d'évaluations fondées sur des méthodes statistiques, en les incitant à se repositionner et à redéfinir les caractéristiques de leur expertise ? Ensuite, comment l'usage des évaluations dans le contrôle et la fabrique de la loi témoignent-elles du rôle alloué par les décideurs politiques aux évaluateurs menant des évaluations statistiques des résultats de l'éducation ? Comment, enfin, l'arrivée d'un nouvel acteur international dans le champ des politiques éducatives, l'OCDE, et de son outil, l'évaluation PISA, a-t-elle contribué à redéfinir les instruments de l'action publique en éducation ? Pour répondre à ces questions, nous observerons ce qu'implique l'arrivée de PISA depuis sa création au milieu des années 1994, jusqu'aux élections présidentielles et sénatoriales de 2017, qui marquent une rupture temporelle politique forte et la fin de la période de recherche de terrain de cette thèse.

L'évaluation des politiques éducatives : d'une prérogative nationale à une logique internationale

L'outil « évaluation » est entré tardivement dans le champ de la politique éducative française, dans les années 1980⁴. Pratiquée depuis plus de 20 ans dans les pays anglo-saxons et scandinaves où elle a été l'outil de légitimation de certaines politiques publiques à la manière des politiques du New Deal dans l'Amérique des années 1960, la démarche d'évaluation a progressivement été adoptée en France sous l'influence de l'émergence à partir des années 1980 du modèle du Nouveau Management Public⁵, appliquant les outils de gestion privés à la gestion publique pour améliorer l'efficacité de l'action publique. A cette

⁴ Clément Lacouette Fougère et Pierre Lascoumes, « L'évaluation : un marronnier de l'action gouvernementale ? », *Revue française d'administration publique*, vol. 4 n° 148, 2013.

⁵ B. Guy Peters, « Nouveau management public (*new public management*) », Laurie Boussaguet éd., *Dictionnaire des politiques publiques. 4^e édition précédée d'un nouvel avant-propos*. Presses de Sciences Po, 2014.

époque, l'évaluation est encore faiblement institutionnalisée⁶ en France, malgré son invocation régulière par les décideurs publics depuis les années 1960, du fait principalement à la grande pluralité des acteurs qui peinent à agir en concert. Aux injonctions d' « *accountability* » auprès des citoyens et de préparation des réformes de façon la plus « efficiente » possible – notamment en réduisant autant que possible la dépense publique – s'ajoute celle de la modernisation de l'État. L'évaluation devient la prérogative de l'État et la condition de son adaptation permanente aux exigences de son temps et de ses administrés.

Elle s'organise autour d'une administration qui met en œuvre les conditions de prise en compte de ce nouveau « souci de soi ⁷ » de l'État, et institue le principe d'évaluation. Appliquée au cas français et aux politiques éducatives, l'institution chargée des évaluations est interne au ministère de l'Éducation nationale et se nomme la Direction de l'évaluation, de la prospective et de la performance. Elle mène l'ensemble des évaluations nationales auprès des élèves, ainsi qu'une série d'évaluation du fonctionnement du système éducatif, toutes fondées sur des méthodes statistiques. Avec la construction d'évaluations statistiques pour évaluer l'action publique, le scientifique, le statisticien, seul à même, selon les préceptes du Nouveau Management Public, de développer les outils pouvant donner une représentation exploitable des résultats aux questions de pilotage, devient expert au sens où il intervient dans une « *situation d'expertise*⁸ », en se positionnant à la « *rencontre d'une conjoncture problématique et d'un savoir spécialisé*⁹ ».

⁶ Par institutionnalisation, nous faisons nôtre ici la définition donnée par Jacques Lagroye, Bastien François et Frédéric Sawicki : ce qui résulte « *de processus qu'on ne saurait réduire à la codification des dispositifs et des règles, à la production de textes législatifs et de règles, à l'adoption d'un nouveau vocable... Elle officialise et donne en modèle des pratiques et des savoir-faire d'échanges entre acteurs, comme par exemple l'utilisation des mêmes catégories par les agents de l'État et les associations* », Jacques Lagroye, Bastien François et Frédéric Sawicki, *Sociologie politique*, Paris, Presses de Sciences-Po – Dalloz, 2002.

⁷ Philippe Bezes, *Réinventer l'État, les réformes de l'administration française (1962-2008)*, Paris, PUF, 2009.

⁸ Yann Bérard et Renaud Crespin, « Situation d'expertise », Emmanuel Henry éd., *Dictionnaire critique de l'expertise. Santé, travail, environnement*. Presses de Sciences Po, 2015.

⁹ Jérôme Péliste, Caroline Protais, Keltoume Larchet, Emmanuel Charrier (dir.), *Des chiffres, des maux et des lettres. Les experts judiciaires en économie, en psychiatrie et en traduction*, Paris, Armand Colin, coll. « Recherches », 2012.

L'évaluation de la « qualité » de l'éducation est une prérogative centrale et ancienne de l'État en France¹⁰, associée étroitement aux grandes réformes républicaines de la fin du XIXe siècle et notamment à l'inscription dans la loi des principes de l'école républicaine : gratuite, laïque et obligatoire. Alors qu'avec la DEPP les autres corps évaluateurs en France sont traditionnellement des corps d'État (les corps d'inspection, ou encore la Cour des Comptes), la fin du XXe siècle voit émerger de nouvelles institutions évaluatrices, y compris dans ce champ des politiques éducatives : les organisations internationales. La production des évaluations, et en particulier celles sur lesquelles portera cette thèse ayant trait aux résultats des acquis des élèves, n'est pas la fin en soi de ceux qui les mènent. Évaluateurs nationaux et internationaux mettent leur expertise au service de la production et de la diffusion de savoirs, qu'ils portent auprès des responsables politiques chargés de faire et de réorienter les politiques publiques. Cette expertise dont ils se disent détenteurs, orientée vers un même objectif de contribuer au pilotage de l'action publique, est définie par chaque acteur dans le cadre d'influence qui est le sien et reconnue par ceux qui les emploient à des fins d'action publique.

L'émergence de l'OCDE en acteur incontournable¹¹ du théâtre de l'évaluation des politiques éducatives pose plusieurs questions. La première concerne les conditions qui ont rendu possible celle-ci : comment comprendre qu'en à peine une décennie le dispositif PISA ait pu faire vaciller l'édifice institutionnel constitué autour de la connaissance du fonctionnement du système éducatif et de ses effets au sein du ministère de l'Éducation nationale et de ses directions ? Une seconde série de questions a trait à la fois à la nature de l'organisation internationale et aux spécificités de la conception et du pilotage de la politique éducative, notamment en France. Pour assurer de sa validité,

¹⁰ Voir par exemple Clément Lacouette Fougère et Pierre Lascoumes, « L'évaluation : un marronnier de l'action gouvernementale ? », *Revue française d'administration publique*, vol. 4 n° 148, 2013 ; Annie Fouquet, « L'évaluation des politiques publiques : État(s) de l'art et controverses », *Revue française d'administration publique*, vol. 4 n° 148, 2013.

¹¹ Sotiria Grek, « Educational Governance and Transnational Experts. Europe and Its Emerging Relationship with the OECD », dans Martin Lawn, Romuald Normand, *Shaping of European Education. Interdisciplinary Approaches*, Abingdon, Routledge, 2015.

l'évaluation se doit d'abord d'être impartiale et les évaluateurs hors de tout conflit d'intérêt. Cette exigence est d'autant plus centrale que l'évaluation contribue très largement à partir des années 1980 à la construction et à la consolidation de la légitimité des pouvoirs publics. Tout l'enjeu pour les décideurs politiques qui participent à l'évaluation est donc de définir précisément le cadre de l'évaluation, ses objectifs, d'en vérifier en amont la méthode, afin de s'assurer de ce qui sera mesuré, de ce qui est susceptible de menacer sa légitimité si les résultats ne démontrent pas l'efficacité des politiques menées ou pire, s'ils concluent à son inefficacité.

Ces premières hypothèses posées, les évaluations internationales sont donc complexes à au moins deux égards. D'une part elles sont pilotées par une instance supranationale qui décide de ce qu'elle veut mesurer. Même si elle associe dans sa prise de décision les pays participants aux évaluations, elle conserve le pouvoir d'orienter les discussions, de réguler des débats et finalement de trancher. D'autre part, du point de vue méthodologique, elle doit prendre en compte la multiplicité des contextes d'évaluation de tous les pays évalués. Impossible pour un pays de déterminer unilatéralement les objectifs, le contexte et la méthode évaluative en choisissant celle qui correspondrait le mieux à ses besoins/attentes en termes d'évaluation de son action. Les administrations nationales remettent donc aux mains d'une organisation internationale les clés d'une évaluation qui peut potentiellement remettre en cause leur légitimité et qui ne sera pas totalement adaptée à leur contexte puisqu'issue du consensus de pays aux systèmes éducatifs différents. Ce risque de délégitimation est d'autant plus fort que les évaluations internationales sont par construction des évaluations externes, conçues par des évaluateurs extérieurs au système éducatif français et à son administration. Au-delà de permettre la certification ou non de l'acquisition de compétences nécessaires pour la poursuite de la scolarité ou pour l'entrée sur le marché du travail, elles ont de plus en plus nourri l'objectif d'évaluer la politique éducative menée par les acteurs du système éducatif.¹²

¹² Lucie Mottier Lopez, « L'évaluation en éducation : des tensions aux controverses. » Dans Lucie Mottier Lopez et M. Crahay (dir.), *Évaluations en tension : entre la régulation des apprentissages*

Dans le cas de PISA, l'évaluation de l'OCDE lancée à la fin des années 1990 sur laquelle portera nos travaux, l'expertise de l'organisation internationale se matérialise certes par une maîtrise approfondie et internationalement reconnue de techniques d'évaluation statistique, mais aussi par une conception nouvelle des objectifs de l'éducation : fournir aux futurs adultes, actifs, citoyens, des *compétences utiles au quotidien* lorsqu'ils seront entrés dans la vie active, plus que des connaissances académiques. Cette légitimité méthodologique et la pratique du consensus dans la définition des objectifs principaux et intermédiaires de son évaluation PISA a inscrit l'OCDE au sein d'une communauté épistémique au sens où l'entend P. Haas : des « *réseaux de professionnels ayant une expertise et une compétence reconnues dans un domaine particulier qui peuvent faire valoir un savoir pertinent sur les politiques publiques du domaine en question*¹³ » Dans le mouvement d'internationalisation des politiques publiques, et les politiques éducatives ne font pas exception, l'OCDE entre pleinement dans le jeu du pilotage des politiques publiques et tente de trouver une place au sein d'une communauté épistémique¹⁴ constituée autour de la production et la diffusion de connaissance et à l'entreprise politique commune : donner à voir les résultats de l'éducation et apporter des éléments de connaissance permettant d'orienter les politiques publiques afin de les améliorer. Cette communauté épistémique, si elle permet de dépasser la conception des organisations internationales comme une somme des arbitrages faits par les gouvernements qui la composent¹⁵, ne désigne pas ici un groupe homogène d'acteurs aux intérêts similaires. Nous verrons que la définition d'un objectif commun et d'un cadre de méthode commun n'efface pas les

et le pilotage des systèmes (pp. 7-25). Bruxelles : De Boeck, 2009; Christian Maroy, Politiques et outils de « l'école de la performance » : accountability, régulation par les résultats et pilotage. Dans C. Maroy (dir.), *L'école à l'épreuve de la performance : les politiques de régulation par les résultats* (pp. 14-31). Bruxelles : De Boeck, 2013.

¹³ Peter Haas, « Introduction: Epistemic Communities and International Policy Coordination », *International Organization* n°1, 1992, p.3.

¹⁴ Peter M. Haas, « Epistemic Communities and International Policy Coordination », *International Organization*, 46/1, 1992 ; Morgan Meyer et Susan Molyneux-Hodgson, « Communautés épistémiques » : une notion utile pour théoriser les collectifs en sciences ? », *Terrains & travaux*, vol. 18, n°1, 2011.

¹⁵ Sandrine Kott, « Une "communauté épistémique" du social ? Experts de l'OIT et internationalisation des politiques sociales dans l'entre-deux guerres », *Genèses*, 2/71, 26-46, 2008.

concurrences entre les différents types de pratiques sur les scènes nationales¹⁶. La dimension politique des connaissances produites¹⁷ et l'imbrication de cadres d'influence exacerbe au contraire des situations de concurrence entre les professionnels de l'évaluation statistique appliquée aux politiques éducatives.

Nous voulons ici montrer que l'irruption de l'évaluation PISA a profondément modifié le champ du pouvoir scolaire français en déplaçant le rôle d'expert des évaluateurs nationaux institutionnels mobilisant des outils statistiques. Nous questionnerons la façon dont l'OCDE est parvenue à entrer et à se faire une place parmi les acteurs clés de l'Éducation nationale en France, malgré sa nature d'organisation internationale.

L'OCDE et la DEPP, inégaux objets de recherche

Si la DEPP a peu fait l'objet de travaux dédiés en science politique, en sociologie des organisations ou encore en sociologie de l'éducation, ce n'est pas le cas de l'OCDE. Ses travaux et son fonctionnement ont été l'objet de nombreuses recherches – certes récentes notamment en France¹⁸ – s'inscrivant dans le champ de l'étude des organisations internationales en elles-mêmes et celui du transfert de politiques publiques et des échanges de connaissances comme forme d'action politique entre les niveaux nationaux et les institutions internationales. Considérée comme « arènes de la gouvernance éducationnelle »¹⁹ et désormais acteur incontournable dans le domaine de l'éducation²⁰, elle a fourni un objet d'étude pour l'observation de la diffusion de « bonnes pratiques » de politique

¹⁶ Sandrine Kott, *art. cit.*

¹⁷ Madeleine Akrich, « From Communities of Practice to Epistemic Communities: Health Mobilizations on the Internet », *Sociological Research Online*, 15/2, 2010.

¹⁸ Vincent Gayon, *L'OCDE au travail. Contribution à une sociologie historique de la « coopération économique internationale » sur le chômage et l'emploi (1970-2010)*, thèse pour le doctorat en science politique sous la direction de Brigitte Gaïti, soutenue le 30 novembre 2010 à l'Université Paris Dauphine, p.6-7.

¹⁹ Kerstin Martens, Alessandra Rusconi et Kathrin Leuze, *New Arenas of Education Governance. The Impact of International Organizations and Markets on Education Policy Making*, Palgrave, 2007.

²⁰ Sotiria Grek, « Educational Governance and Transnational Experts. Europe and Its Emerging Relationship with the OECD », dans Martin Lawn, Romuald Normand, *Shaping of European Education. Interdisciplinary Approaches*, Abingdon, Routledge, 2015.

publique par une organisation dépourvue de pouvoir coercitif, non productrice de normes ou de textes législatifs, et ne disposant pas de capacité financière. Ses marges de manœuvre relevant du *soft power* lui ont conféré un intérêt pour travailler les formes et les moyens de pression par les pairs²¹. Organisation prolix, son pouvoir d'influence a été étudié sous le prisme de ses productions et notamment de ses rapports, à la fois en tant qu'outil de transmission de ses orientations en matière de politique publique et en tant que concrétisation tangible d'une fabrique du consensus entre ses pays membres²².

« *L'évaluation ne se réduit pas à un outil, même fondamental, d'une gestion publique repensée, à un étalon de valeur d'un nouveau marché scolaire en expansion, ou à l'avatar d'une planification technocratique reconvertie* », nous dit Xavier Pons en conclusion de son ouvrage sur les évaluateurs de l'action éducative. C'est bien en suivant cette assertion que nous souhaitons travailler sur les effets de recomposition du rôle et des caractéristiques de l'expertise des évaluateurs statistiques provoqués par l'arrivée de PISA parmi les outils d'évaluation de l'action éducative nationale. L'évaluation PISA elle-même a suscité de nombreux travaux, en France comme dans d'autres pays participants à l'évaluation²³. Ils ont abordé l'évaluation principalement du point-de-vue de l'usage qui en est fait ou qui n'en est pas fait, ou encore du point-de-vue des réactions que ses résultats ont suscité dans les discours aussi bien politiques que médiatiques. L'effet de l'outil d'une organisation internationale à la légitimité forte dans le champ de la politique économique sur les stratégies d'influence et de définition des rôles des acteurs du champ de la politique éducative n'a en revanche pas connu de développements majeurs. C'est à cette analyse que nous

²¹ Dorota Dakowska, « What(ever) works : les organisations internationales et les usages de « bonnes pratiques » dans le domaine de l'enseignement supérieur », *Critique Internationale*, Presses de sciences po, 2017.

²² Vincent Gayon, *L'OCDE au travail...*, thèse cit.

²³ Xavier Pons et Nathalie Mons ont largement contribué à créer et à nourrir les travaux de recherche universitaire sur PISA. Une partie de leurs travaux ont servi à construire le cadre d'analyse de cette thèse et sont donc cités en bibliographie et au cours du texte. Ils ont par ailleurs participé à un projet de recherche européen Know&Pol, qui a amené des chercheurs de plusieurs pays de l'Union Européenne, participants à PISA, à travailler sur la réception de ses résultats notamment par les parlements nationaux. Ces travaux, portant sur la France mais aussi sur le Portugal et l'Irlande notamment, sont également cités en bibliographie en fin de volume.

voulons contribuer par cette thèse. Du côté des évaluations menées par la DEPP, quelques travaux ont traité spécifiquement de l'évaluation CEDRE au prisme de ses liens avec les évaluations internationales et notamment PISA et TIMMS²⁴. Le reste de la littérature concernant la DEPP est constituée principalement des travaux de Xavier Pons sur les évaluateurs de l'action éducative en générale²⁵, la DEPP en faisant partie parmi d'autres institutions.

Nous proposons ici de compléter ces travaux en regardant la façon dont l'entrée de l'OCDE dans le champ de l'évaluation des politiques éducatives nationales a induit une redéfinition – voire une remise en question – du rôle et de la nature de l'expertise des agents de la DEPP. Cette redéfinition reflète d'une part les stratégies des agents de l'évaluation nationale eux-mêmes qui cherchent à valoriser leurs travaux et leur maîtrise technique des évaluations statistiques auprès de ceux qui en sont les destinataires premiers : les décideurs politiques nationaux. D'autre part, elle est induite par les usages différenciés que ces décideurs politiques font des évaluations mises à leur disposition, à la fois dans leur action quotidienne de contrôle de l'action du gouvernement en la matière, soit aux moments de réformes majeures du système éducatifs.

Nous verrons que le champ des acteurs de l'évaluation n'est pas statique, mais qu'il se recompose en fonction à la fois du contexte politique dans lequel les évaluations sont produites et publiées, mais aussi en fonction de ce que permettent les évaluations au fur et à mesure du temps, et notamment des comparaisons temporelles entre plusieurs cycles évaluant les mêmes compétences à quelques années d'intervalle. Nous observerons deux grands types d'interactions. D'abord, nous verrons comment les interactions des deux organisations, OCDE et DEPP, et des acteurs qui les composent, lors de la construction et de la conduite de PISA, ont un effet sur la définition des rôles d'experts de chacun. Ensuite, les évaluations menées pour rendre compte des

²⁴ Elodie Bellarbre, *Compenser les inégalités dans le secondaire en France : études à partir des enquêtes PISA, CEDRE et TALIS*, thèse pour le doctorat d'économie sous la direction de Jean-François Giret, soutenue à l'Université de Bourgogne Franche Comté le 19 octobre 2017.

²⁵ Xavier Pons, *Evaluer l'action éducative, des professionnels en concurrence*, Education et société, PUF, Paris, 2010.

résultats de l'éducation sont conçues pour aider au pilotage de l'action éducative. Nous analyserons en quoi les stratégies de définition des expertises de l'OCDE d'une part et de la DEPP d'autre part, ont porté leur fruit en étant convoquées par les parlementaires dans la fabrication de la loi et le contrôle de l'action de l'État. En reprenant à notre compte les travaux sur les conditions paradigmatiques de changement des politiques publiques de Peter Hall²⁶, nous poserons l'hypothèse que tout transfert d'outils d'évaluation nécessite un référentiel particulier, spécifique à la façon d'envisager non seulement l'objet de l'évaluation mise en œuvre mais aussi les vecteurs choisis pour cette mise en œuvre et les objectifs à atteindre. Le référentiel doit être « *compréhensible et plausible par les acteurs en jeu* ²⁷» et définit le champ des possibles commun entre les acteurs ou les organisations. Afin de saisir la façon dont les référentiels ont pu évoluer et conférer une légitimité différente d'une période à une autre aux évaluations statistiques des résultats de l'action éducative, y compris PISA, nous resserrerons la focale autour de moments critiques qui mettent en lumière les logiques nouvelles de pouvoir et d'influence qui émergent parmi les acteurs de la politique éducative. Ils permettront d'observer en particulier les dynamiques relatives aux acteurs directement investis dans les évaluations nationales, soit au plan technique avec la DEPP, soit au plan politique avec les gouvernements successifs et les parlementaires qui contrôlent leur action.

²⁶ Peter Hall distingue trois formes de changement des politiques publiques, de trois ordres. Le changement de premier ordre est une modification du niveau d'utilisation des instruments. Le second correspond à l'introduction de nouveaux instruments de politiques publiques (et donc à une nouvelle conception de l'objet à évaluer). Le changement de troisième ordre est un changement simultané des trois dimensions des politiques publiques données par Hall : leur but, leurs instruments et leur mode d'utilisation. Voir Peter Hall, « Policy Paradigms, Social Learning, and the State : The Case of Economic Policymaking in Britain », *Comparative Politics*, Vol.25, N°3, 1993.

²⁷ James E. Anderson, *Public Policymaking*, Praeger, 1975. Voir aussi les travaux de Pierre Muller mobilisant ce concept de référentiel, Pierre Muller « L'analyse cognitive des politiques publiques : vers une sociologie politique de l'action publique », *Revue française de science politique*, 50^e année, n°2, 2000.

L'évaluation, instrument d'action publique et vecteur de pouvoir ?

Cette thèse fait l'hypothèse que l'évolution des pratiques d'évaluation est un processus susceptible de déplacer les rapports de force du champ du pouvoir scolaire au sens où l'entend Pierre Clément, c'est-à-dire « *l'espace des positions sociales à partir desquelles peut s'exercer un pouvoir sur « l'école »* ». En s'intéressant tout particulièrement à la naissance, à la construction, et à l'organisation de l'évaluation PISA par l'OCDE, nous ambitionnons de porter un regard compréhensif sur les effets, sur l'organisation des relations entre certains acteurs du champ du pouvoir scolaire français, de ces évaluations internationales en éducation menées par des évaluateurs internationaux en éducation. Notre travail s'inscrit en ce sens aussi dans le courant d'étude des instruments d'action publique, conçus par Lacoumes et Le Gales comme un « *dispositif à la fois technique et social qui organise des rapports sociaux spécifiques entre la puissance publique et ses destinataires en fonction des représentations et des significations dont il est porteur* ». Les instruments ne sont plus alors les paramètres explicatifs de l'action publique mais à l'inverse l'action publique peut être déconstruite par le biais de ses instruments, qui en constituent la structure.

Cela amènera à s'interroger sur l'instrument PISA par rapport aux « instruments évaluation » de la DEPP. Si tous répondent à la dénomination « évaluation », et au-delà des diversités méthodologiques qui les caractérisent, sont-ils réellement les mêmes instruments pour l'action publique ? L'usage croissant des outils statistiques depuis les années 1980 et la création d'évaluations standardisées en éducation témoignent d'une évolution importante de la façon de piloter l'action publique. Le rôle des évaluateurs et des méthodes d'évaluations est un prisme d'analyse intéressant pour mettre en évidence la façon dont chaque évaluateur se positionne dans le champ de l'action publique en éducation. La méthodologie des évaluations menées à la fois par la DEPP et par l'OCDE pour PISA a été un facteur de différenciation structurant au départ. Nous montrerons ainsi comment la méthode devient, au même titre que l'objectif de l'évaluation, l'objet de transferts entre l'OCDE et la DEPP et comment cela permet aux agents

de redéfinir leur rôle dans le cycle de vie des politiques éducatives nationales. La diffusion recherchée de « bonnes pratiques » de politiques publiques se double avec PISA d'une diffusion de « bonnes pratiques » d'évaluation des politiques éducatives. Si elle n'est pas officiellement revendiquée par l'organisation qui ne cherche pas à imposer son modèle d'outil aux pays membres, elle transparaît néanmoins en ce qu'elle est considérée comme représentative du consensus des participants à PISA la reconnaissant la mieux à même de rendre compte des objectifs et des résultats de l'éducation dans un monde devenu à leurs yeux incertain. Nous verrons en quoi cette communication autour de la « bonne façon d'évaluer les résultats de l'éducation », dont l'effet est renforcé par l'adhésion progressive des responsables politiques des pays participants à PISA et notamment des dirigeants français, a contribué à mettre au second plan l'expertise récemment institutionnalisée des évaluateurs de la DEPP.

Si l'évaluation est devenue un instrument de pilotage des politiques publiques, ce n'est pas simplement car elle est susceptible d'apporter, à la commande des décideurs politiques, des éléments chiffrés permettant de justifier du bienfondé de leur action. Elle est aussi inséparablement devenue l'instrument de l'identification, de la popularisation et de la mise à l'agenda des problèmes publics²⁸, c'est-à-dire des problèmes sociaux sur lesquels il convient d'agir. Progressivement intégrée de façon homogène dans les pratiques des élus de l'ensemble des familles politiques, elle devient en outre un outil du contrôle de l'action de l'État et de la fabrique de la loi. Nous verrons que PISA en sera un catalyseur. Et par ce statut que confère l'évaluation à son producteur, elle devient porteuse du pouvoir d'inférer sur la réorientation de l'action éducative.

L'étude des acteurs du champ de l'évaluation des politiques éducatives est relativement peu présente dans les travaux de sociologie et de science politiques. Cette thèse se veut contribuer à nourrir ce domaine de recherche. Xavier Pons est sans doute à l'origine des travaux les plus fournis sur le sujet, menés depuis plus de 10 ans sur les évaluateurs de l'action éducative en France, et en particulier les

²⁸ Erik Neveu, *Sociologie politique des problèmes publics*, Armand Colin, 2015.

acteurs nationaux de l'évaluation, qu'ils aient une mission officielle d'évaluation ou qu'ils participent de façon latente à celle-ci²⁹. L'OCDE fait partie de cette dernière catégorie et est étudiée comme telle par Pons, spécifiquement parce qu'elle a conçu et mène son évaluation PISA depuis la fin des années 1990. En association avec Nathalie Mons, auteure elle aussi de nombreuses publications portant sur PISA, il et elle interrogent la façon dont ministres, parlementaires et médias ont pu recevoir les résultats de l'évaluation et les utiliser pour leurs propres travaux. Les acteurs de la DEPP sont présents dans leurs travaux puisqu'ils sont professionnels de l'évaluation en éducation et impliqués dans PISA. Néanmoins, la recomposition de leur rôle et des caractéristiques de leur expertise dès le début de leur travail sur PISA, en miroir de ce que les acteurs de l'OCDE démontrent de leur propre expertise, reste peu étudiée. C'est cette zone d'ombre que la présente thèse ambitionne d'éclairer.

Deux groupes d'acteurs seront au centre de ces travaux. Ceux que nous appellerons les « décideurs politiques » d'abord, c'est-à-dire les acteurs de l'exécutif qui déterminent les orientations et les sujets de la politique éducative française, et les élus appartenant au pouvoir législatif, qui contrôlent l'action gouvernementale et contribuent à voter et contribuer aux réformes. Ce sont principalement les différents ministres de l'éducation et les cabinets qui transmettent les décisions à ceux en charge de définir leurs possibilités et leurs modalités d'application, ainsi que les parlementaires députés et sénateurs.

Ceux que nous appellerons les « évaluateurs » ensuite, qui sont chargés de mener les évaluations quantitatives des résultats de la politique éducative conduite par les gouvernements successifs. Ils sont nationaux, les agents de la DEPP, et internationaux, les agents de la Direction de l'Éducation de l'OCDE. En suivant l'approche proposée par Xavier Pons dans ses travaux pionniers sur les professionnels de l'évaluation de l'action éducative en France, nous avons choisi d'identifier les évaluateurs statistiques de l'action éducative comme étant ceux qui se considèrent comme évaluateurs de l'action éducative. Cette définition

²⁹ Xavier Pons, *Evaluer l'action éducative, des professionnels en concurrence*, Education & Société, PUF, Paris, 2010.

a l'avantage de détacher nos travaux d'une démarche d'étude trop descriptive des acteurs, qui ne permettrait pas de mettre en évidence leur (re)définition de leur rôle et de l'expertise sur la période qui nous intéresse. Prendre comme objet les évaluateurs parce qu'ils se définissent comme évaluateur de l'action éducative donne à voir la façon dont les expertises d'entremêlent et dont les zones d'influence identifiées par les acteurs permettent d'en tirer ou non profit pour légitimer sa propre action. Si ces évaluateurs produisent tous des indicateurs statistiques et des évaluations standardisées, ont-ils structuré leurs zones d'influence autour de la réalisation d'évaluations de la même façon ? Comment l'entrée en jeu de l'OCDE avec son PISA, qui contribue au déplacement vers la scène internationale de l' « *arène de la gouvernance éducationnelle*³⁰ » implique t'elle pour la DEPP de réaffirmer sa légitimité d'expert des résultats de l'éducation ? Et enfin, les stratégies de reconquête du pouvoir d'expert des agents de la DEPP, au sens où ils peuvent contribuer à justifier ou déterminer la réorientation de l'action publique en éducation, ont-elles réussi ? Pour répondre à ces questions, nous étudierons les interactions mouvantes entre ces acteurs de la même communauté épistémique, entre concurrence manifeste et logiques de valorisation de leurs travaux auprès de destinataires politiques ou académiques, nationaux ou internationaux.

L'expertise institutionnelle en politique éducative : des savants politiques ?

Les travaux de recherche relatifs à l'expertise et à ses acteurs n'ont cessé de gagner en ampleur à partir des années 1990. En particulier, le concept d'expertise et son importance dans le champ de l'aide à la décision de politique publique a été l'objet de plusieurs courants de recherche à la fois dans le champ des sciences juridiques, de la sociologie des acteurs et des sciences politiques³¹. Certains travaux ont étudié les manifestations et l'origine de l'expertise allouée à

³⁰ Kerstin Martens, Alessandra Rusconi et Kathrin Leuze, *New Arenas of Education Governance. The Impact of International Organizations and Markets on Education Policy Making*, Palgrave, 2007.

³¹ Voir au chapitre 3 une revue de la littérature plus fournie sur l'expertise et les experts sont développés.

un expert³², allant des compétences revendiquées acquises par l'expérience faisant de l'expert un expert en tout temps, à l'expertise attribuée à certains groupes professionnels par défaut³³. D'autres approches ont privilégié l'appréhension de l'expertise à travers la commande politique et la mobilisation de savoirs experts. Elles examinent l'émergence de situations d'expertises, de moments de conjonction entre le besoin de résoudre une situation problématique en contexte d'incertitude et le savoir de celui qui peut alors y contribuer et devient alors expert³⁴. Elles s'intéressent également à l'instrumentalisation politique de l'expertise accordant à l'indépendance de l'expert une importance prépondérante pour que lui soit attribué ce statut. Des travaux plus récents et plus éloignés de notre cadre conceptuel de recherche s'intéressent à l'expertise et à la contre-expertise profane, certains cherchant à établir le lien entre ces nouvelles formes légitimes d'expertise et la crise de légitimité de l'expertise notamment scientifique³⁵. Enfin, dans le sillage de Philippe Bezes et de ses travaux sur le souci de soi de l'État, l'expertise institutionnelle a constitué un objet de recherche relativement récent permettant d'aborder à la fois son effet sur la reconfiguration de l'action publique, l'intégration de modes d'évaluation issus du monde managérial dans les pratiques d'expertise d'État³⁶ et le repositionnement de la frontière entre savant et politique à travers l'émergence d'experts au rôle politique fort, y compris internationaux³⁷.

³² Jean-Yves Trépos, *La sociologie de l'expertise*, PUF, Que-sais-je ?, 1996.

³³ Andrew Abbott, *The System of Professions, an Essay on the Division of Expert Labor*, University of Chicago Press, 1988 ; Jean-Yves Trépos, op. cit. ; Claude Dubar, Pierre Tripier et Valérie Boussard, *Sociologie des professions*, Paris, Armand Collin, 2011.

³⁴ Voir par exemple les Actes de la table ronde organisée par le CRESAL (Centre de Recherches et d'Etudes Sociologiques Appliquées de la Loire) à Saint Etienne les 14 et 15 mars 1985 ; Jacques Chevallier « L'entrée en expertise », *Politix*, vol.9, n°36, quatrième trimestre, 1996 ; Laurence Dumoulin, Stéphane La Branche, Cécile Robert et Philippe Warin, *Le recours aux experts : raisons et usages politiques*, Presses universitaires de Grenoble, Grenoble, 2005 ; Corinne Delmas, *Sociologie politique de l'expertise*, Paris, La Découverte, 2011 ; Emmanuel Henry et al., *Dictionnaire critique de l'expertise. Santé, travail, environnement*. Presses de Sciences Po, 2015.

³⁵ Voir par exemple Mark Rutgers, Maarten A. Mentzel, « Scientific expertise and public policy : resolving paradoxes ? », *Science and Public Policy*, 26(3), 1999.

³⁶ Corinne Delmas, op. cit. ; Isabelle Berrebi-Hoffmann et Pierre Grémion, « A quoi servent les experts ? », *Cahiers internationaux de sociologie*, n°126, 2009.

³⁷ Les organisations internationales par exemple, et en particulier Vincent Gayon, *L'OCDE au travail...., thèse cit.*

Si l'on se rapporte à l'évaluation, l'étude de ses acteurs et des processus liant évaluation et prise de décision publique a souvent choisi l'angle d'approche de l'expertise et de la production de la connaissance pour étudier les interactions entre les experts et les décideurs publics. La production de connaissance, historiquement le fait de l'université, s'est progressivement étendue à d'autres organismes et institutions, publics ou privés, que ce soit les agences publiques et parapubliques, *think tanks* ou cabinets de conseil en action publique qui produisent des études à la fois d'impact et de résultat pour les gouvernements. Le rôle de la production de la connaissance en politique publique a évolué parallèlement à la multiplication des acteurs et est passé d'une fonction descriptive et analytique à celle d'aide à la décision à toutes les étapes du « cycle de vie » d'une politique publique. Désormais au croisement entre connaissance scientifique et connaissance appliquée, la production des connaissances est pour certains auteurs un outil de gouvernement en soi qui oriente les politiques publiques³⁸. Dans le sillage des travaux en sociologie de la connaissance, la question des enjeux politiques, économiques et scientifiques de la production de la connaissance par les experts nous paraît centrale pour comprendre la production des politiques publiques, en particulier dans le secteur des politiques éducatives³⁹.

Nous nous inscrivons ici dans la lignée des travaux qui s'intéressent aux conditions de travail et aux processus de production des travaux des services statistiques ministériels⁴⁰ et des organisations internationales⁴¹, en analysant la façon dont la production des évaluations nationales et internationales en

³⁸ Jenny Ozga, *Governing Knowledge : Research Steering and Research Quality*, *European Educational Research Journal*, 9(3), Septembre 2008.

³⁹ Romuald Normand, *Gouverner la réussite scolaire. Une arithmétique politique des inégalités*. Berne : Peter Lang et Lyon ENS de Lyon, 2011.

⁴⁰ Etienne Pénissat, sur la DARES et notamment ses travaux de thèse : *l'Etat des chiffres : sociologie du service de statistique et des statisticiens du ministère du Travail et de l'Emploi (1954-2008)*, soutenue en 2009 à Paris, EHESS, sous la direction de Michel Offerlé ; Nicolas Belorgey, « Les cadres de l'expertise publique. Les trois dimensions de l'expertise d'État saisies par une ethnographie des économistes de la Banque de France », *Genèses*, 85, 2011, sur la publication des études économiques de la Banque de France.

⁴¹ Voir les travaux de Vincent Gayon, sur l'OCDE et les rapports économiques, en plus de sa thèse déjà citée, Vincent Gayon, « Un atelier d'écriture internationale : l'OCDE au travail. Éléments de sociologie de la forme « rapport » », *Sociologie du travail*, Vol. 51, n° 3, 2009.

éducation confère une expertise à leurs producteurs. Nous prendrons pour définition initiale de l'expertise celle fournie par Philippe Roqueplo au début des années 1990 : « *L'expression d'une connaissance scientifique ne revêt valeur d'expertise que dans la mesure où elle s'articule à un processus décisionnel et c'est précisément cette articulation qui lui confère sa valeur d'expertise.* »⁴² On voit ici poindre l'un des effets possibles de l'utilisation avant toute autre évaluation de PISA par les décideurs politiques. Si les évaluations nationales ne sont plus employées à des fins d'action publique, ceux qui les produisent sont-ils toujours des experts ? Nous verrons en quoi cette remise en question de la légitimité politique des évaluations nationales a pu inciter les évaluateurs nationaux, en l'occurrence les agents de la DEPP, à tenter de reconfigurer leur rôle pour retrouver cette prérogative centrale de leur identité, quitte à accepter de devenir des « passeurs⁴³ », diffuseurs des résultats (concurrents) de PISA auprès des décideurs politiques nationaux.

« le PISA » ou « PISA » ? De l'acronyme au nom propre

Il est frappant de constater à la lecture des travaux sur l'évaluation de l'OCDE, et des prises de parole des agents de la DEPP et des acteurs politiques français, qu'ils soient membres de l'exécutif ou législateurs, que le Programme international pour le suivi des acquis des élèves, PISA, est passé d'un acronyme de nom commun à un nom propre. On ne parle pas « du » PISA mais bien « de » PISA, lui attribuant par cette formulation un pouvoir d'attraction ou de rejet similaire à celui que peut porter en elle une marque auprès des consommateurs. Le « produit » PISA est d'ailleurs défini comme tel comme cela par les interlocuteurs que nous avons pu rencontrer à l'OCDE dans le cadre d'entretiens

⁴² Philippe Roqueplo, « L'expertise scientifique, consensus ou conflit ? », *La terre outragée, les experts sont formels*, Autrement, série sciences en société, n°1, 1992, p.162.

⁴³ Fabien Jobard et al. « Sociologie politique des passeurs. Acteurs dans la circulation des savoirs, des normes et des politiques publiques », *Revue française de science politique*, vol. 70, n°5, 2020.

individuels ou de discussions informelles. PISA est « brandé ⁴⁴», il fait l'objet d'une stratégie sciemment orchestrée de gestion d'une image de marque et de lancement de produits dérivés⁴⁵. Le « produit » PISA a été travaillé par l'OCDE pour porter en lui des qualités de foi⁴⁶, « *credence qualities* » en anglais, qui doivent être crues sans pouvoir être prouvées par l'usage et dont les utilisateurs et les participants à PISA sont convaincus⁴⁷. Cette adhésion à (au) PISA, ou plus exactement à la légitimité des résultats fournis par (le) PISA et des conclusions formulées par l'OCDE à partir d'eux, est un axe d'analyse central de cette thèse.

PISA est bien un produit aux yeux des acteurs, et nous avons décidé dans cette thèse de conserver la dénomination par nom propre et non par acronyme. Nous parlerons ici de PISA et non pas du PISA, pour comprendre l'évaluation internationale comme un objet, une entité porteuse de principes de fonctionnement et de représentations propres, allant au-delà du sens des mots qui les composent. Cela n'implique en rien l'adhésion à la marque PISA, mais c'est une façon de reconnaître la pratique des acteurs dans le cadre d'une évaluation à laquelle ils accordent un statut particulier. De la même façon, dans un souci d'équilibre des axes d'analyse des effets de PISA sur la légitimité de l'expertise des évaluateurs nationaux, nous nous référerons de cette même manière aux autres évaluations qui seront mobilisées ici. Nous parlerons donc de CEDRE⁴⁸, de TIMMS⁴⁹, de PIRLS⁵⁰, ou encore de IALS⁵¹.

⁴⁴ De l'anglais « brand », qui signifie « marque ». Le branding est une sous-activité du marketing qui a émergé dans la seconde partie du XXe siècle et qui consiste à gérer l'image de marque d'un produit, d'une marque ou d'une entreprise.

⁴⁵ De nombreuses publications et évaluations complètes ont été progressivement dérivées de PISA. C'est le cas par exemple des *country notes* qui sont réalisées sur des pays spécifiques, ou des projets *PISA for schools*, ou *PISA for development*.

⁴⁶ Voir sur ce sujet par exemple les classifications des économistes Phillip Nelson, « Information and Consumer Behaviour », *Journal of Political Economy*, vol. 78, 1970 ; Michael Darby et Edi Karni, « Free competition and the optimal amount of fraud », *Journal of Law and Economics*, vol. 16, 1973.

⁴⁷ Nous y reviendrons plus en détail dans le chapitre 2.

⁴⁸ Le Cycle des Evaluations Disciplinaires Réalisées sur Echantillon.

⁴⁹ Trend in Mathematics and Science Study.

⁵⁰ Progress in International Reading Literacy.

⁵¹ International Adult Literacy Survey.

Méthode de l'enquête

Afin d'apporter des réponses à ces questionnements, cette thèse s'appuie sur le croisement de plusieurs méthodes relevant de l'enquête sociologique et de travaux documentaires sur un certain nombre d'archives et de produits de la recherche universitaire sur le sujet.

Entre activité professionnelle et observation participante

La première d'entre elle, à la fois temporellement et pour ce qu'elle a apporté à la problématisation et à la recherche de matériau pour ces travaux, aura sans doute été l'observation participante des contextes de travail de la Direction de l'éducation de l'OCDE, de la DEPP et des assemblées parlementaires. Cette thèse a été menée en parallèle d'activités professionnelles successives et menées à plein temps⁵². J'ai passé quatre mois au sein de la direction de l'éducation de l'OCDE en 2011, pour travailler sur la publication descriptive de l'état des écoles de l'OCDE, Regards sur l'Education. L'évaluation PISA n'avait pas alors encore le statut qui est le sien aujourd'hui. La publication de ses résultats bénéficiait d'une communication du même niveau que d'autres rapports annuels très attendus comme Regards sur l'Education. Si ce passage à « EDU », comme est appelée la direction par ses agents, n'a pas été directement en lien avec PISA, elle a néanmoins permis d'observer la façon dont les équipes de la direction travaillent, dont elles échangent avec les représentants des pays membres. Le management de l'OCDE étant à la fois très vertical dans la façon dont fonctionnent les équipes « projet », et horizontal dans le souci permanente de construire un esprit d'équipe au-delà des grades et postes de chacun. Cette mission m'aura permis de rencontrer certains des acteurs clés, qui ont ensuite

⁵² En quelques mots, j'ai travaillé en 2011 au sein de la Direction de l'éducation de l'OCDE pour la publication Regards sur l'Education, puis en 2012 pour la DEPP à l'évaluation d'un volet du Livret de compétences expérimental prévu dans la loi de novembre 2009 relative à l'orientation et à la formation professionnelle tout au long de la vie. J'ai ensuite occupé le poste de collaboratrice parlementaire en charge des affaires sociales au Sénat pour les parlementaires du groupe écologiste de 2013 à la fin d'année 2017.

immédiatement accepté de me rencontrer en entretien lorsqu'ils ont été sollicités en ce sens. Un passage court à l'UNESCO par la suite, dans le cadre d'un autre projet toujours sur l'évaluation des résultats de l'éducation, m'aura offert l'occasion en comparaison d'isoler des postures spécifiques des agents de l'OCDE, des manières de travailler propres à l'OCDE, qui auront été éclairantes au moment de définir les traits de l'expertise des agents travaillant à la production de PISA.

Trois mois auront été consacrés ensuite à un projet d'évaluation au sein de la DEPP, donnant à voir de l'intérieur l'organisation et le fonctionnement des équipes mais aussi les liens avec la tutelle ministérielle. Ces temps courts n'ont pas eu pour effet de créer un lien de grande proximité avec les acteurs en poste dans chaque institution, éloignant le risque de partialité du regard du chercheur. En revanche, ils ont fourni un cadrage éclairant sur le fonctionnement de chaque institution, et m'ont permis encore une fois de créer des contacts et des liens qui faciliteront ensuite les sollicitations concernant des entretiens individuels approfondis

L'analyse des stratégies du groupe à l'aune des acteurs

L'observation du fonctionnement des organisations sur lesquelles porte ce travail et des relations en interne entre les différents protagonistes, évaluateurs statistiques et décideurs politiques, a été complétée d'une série de 23 entretiens semi-directifs. Ils ont été menés principalement auprès de responsables de la DEPP en charge de suivre le déploiement et les résultats de PISA pour la France, et de représenter la France auprès de l'OCDE, avec des membres de l'équipe PISA-OCDE et en particulier les membres de la première équipe fondatrice de l'évaluation, des parlementaires et leurs équipes, des membres de cabinet ministériel. Ils ont été nourris et enrichis par les discussions informelles et les réunions de travail auxquelles j'ai pu participer pendant plusieurs années. Les entretiens ont duré souvent plus d'une heure, et ont eu lieu quand cela était possible dans les bureaux des personnes. Certains ont été rencontrés deux fois en

entretien formel, permettant de reposer parfois certaines questions, pour s'assurer de la justesse de compréhension des enchaînements d'événements ou de l'uniformité du discours sur les moments de tension ou de succès. Ces entretiens semi-directifs ont été indispensables à au moins deux égards.

D'abord, ils traduisent le regard de terrain indispensable pour incarner, mettre en perspective les informations recueillies et leur donner sens dans ces travaux. Ils ont permis de répondre aux questions soulevées par des situations d'observation ou des consultations de documents d'archives ou de travaux de recherche ayant donné des éléments d'information demandant à être recontextualisés et perçus à la lumière des jeux d'acteurs en cours à l'époque.

Ensuite, ils ont apporté des éléments précieux pour retracer, du point-de-vue descriptif, les relations qu'ont entretenues les agents de la DEPP avec les équipes de l'OCDE dès sa construction, et saisir comment les relations de travail ont évolué au cours du temps. La plupart des documents d'archives sur cette période sont encore classifiés et la plupart des demandes de déclassification que j'ai formulées ont été rejeté. Il était alors particulièrement compliqué de saisir avec précision les problématiques politiques et méthodologiques posées par les différentes phases de construction de l'évaluation PISA, les temps de réunion et de concertation, le rôle attribué par l'OCDE aux agents de la DEPP – comme tout autre opérateur national des pays participants à PISA d'ailleurs – dans la préparation des cycles avec leurs différentes étapes, et les marges de manœuvre des agents ministériels dans la formalisation des indicateurs finaux.

Des archives officielles complétées de littérature « grise »

Ce travail s'appuie ensuite sur une série de documents d'archives de l'OCDE sur la période 1994-2019, date à laquelle est publié le premier rapport portant sur le 7^e cycle PISA. Ces archives, fournies par le centre des archives de l'OCDE pour une grande partie d'entre elles et accessibles en ligne pour une petite partie, ont permis de retracer les étapes clés de la construction de PISA dans les années 1990, en montrant l'évaluation en train d'être conçue. Si les

documents de l'OCDE s'attachent à montrer le consensus en train de se faire, en réduisant souvent au minimum les mentions nominatives des oppositions exprimées lors des réunions de travail ou de pilotage, ils ont néanmoins permis de révéler les questions posées en amont de chaque temps de travail et des décisions prises à leur issue. Ils ont également permis, entre les lignes, à la lecture des rapports spéciaux commandés par le Pisa Governing Board regroupant les pays participants à l'évaluation, de mettre en lumière les points de questionnement ou de tension qui ont pu émerger au fur et à mesure de la construction du projet. Ces sujets ont été ensuite évoqués avec les acteurs rencontrés en entretien et ont permis d'évoquer avec eux les moments de tension ou au contraire les points d'accord sur les objectifs et la méthodologie de PISA.

Ces archives ont été complétées de toute une série de documents constituant une littérature grise, qui a été portée à notre connaissance ou sollicitée auprès des personnes rencontrées en entretien, ou rencontrées dans le cadre de mes activités professionnelles : notes internes, revues de presse réalisées à la sortie des publications PISA, projets de rapports, etc.

Quelques matériaux prosopographiques ont par ailleurs été employés lorsque cela permettait d'éclairer le positionnement des acteurs, leur mobilisation des évaluations allant parfois à l'encontre de ce que leur fonction ou leur appartenance politique aurait pu laisser a priori imaginer. Il a généralement été reconstitué à partir des entretiens réalisés soit avec les acteurs eux-mêmes soit avec des proches ou membres de la même administration, à partir de notices disponibles en ligne ou de travaux de recherche ayant déjà retracé leurs parcours.

De nombreux ouvrages et articles portant sur l'évaluation des politiques publiques, les politiques éducatives, les relations entre acteurs d'organisations, l'évaluation PISA en elle-même, l'expertise et l'action publique, entre autres sujets en lien avec ces présents travaux de recherche auront bien sûr été abondamment parcouru tout au long de ces années de doctorat. Les principales sources universitaires et de recherche utilisées sont citées en bibliographie.

Le journal officiel comme témoin des usages de l'évaluation dans la fabrique de la loi

L'étude de ces archives a été enfin complétée par l'analyse des interventions parlementaires en séance publique sur la période allant de 2000 à 2019, soit de l'année de démarrage de PISA et du premier cycle passé en France (avant même les premiers résultats qui seront publiés en 2001), jusqu'à la publication des résultats principaux de PISA 2018. En particulier, trois types de matériaux auront été plus particulièrement analysés. Ils auront permis de rendre compte des usages différenciés de PISA et des évaluations de la DEPP à la fois en continu dans le cadre des missions de contrôle parlementaire de l'action du gouvernement, et à la fois lors de grands moments de réforme lorsqu'il s'est agi de revoir le fonctionnement du système éducatif à la lumière de ses résultats. Pour nourrir le premier questionnement, les matériaux choisis ont été ceux permettant de donner à voir les problèmes publics jugés urgents par les parlementaires et impliquant une réponse rapide du gouvernement ou nécessitant un travail d'enquête pour apporter des solutions de fond. Les archives et les comptes-rendus des questions au gouvernement, aussi bien écrites qu'orales, aussi bien à l'Assemblée nationale qu'au Sénat⁵³ ont donc constitué une source de premier choix, ainsi que les rapports parlementaires, missions d'information ou commission d'enquête, qui ont mobilisé nos évaluations nationales et/ou internationales. Pour le second, les comptes-rendus détaillés publiés au journal officiel des débats parlementaires en séance publique et des auditions des ministres lors de l'examen des deux grands textes de loi portant sur la refonte du système éducatif français, en 2005 et en 2013, ont fourni des éléments explicatifs d'une grande richesse, repris au chapitre 6.

⁵³ Pour retrouver le détail des matériaux examinés, voir le chapitre 5.

Plan de la thèse

Cette thèse se compose de trois parties comportant chacune deux chapitres.

La première partie dessine le cadre dans lequel les évaluateurs nationaux et internationaux vont entrer en interaction à la fin des années 1990. Elle donne notamment à voir la façon dont l'institutionnalisation tardive et difficile de la pratique de l'évaluation comme outil de pilotage de l'action éducative en France, et donc la constitution d'évaluateurs officiels de l'action de l'État en éducation, coïncide temporellement avec les premières réflexions menées par l'OCDE sur le lancement de son évaluation internationale sur les résultats de l'éducation. Le premier chapitre retrace les étapes menant à la répartition, au début du XXI^e siècle, des missions d'évaluation de l'action publique en éducation entre plusieurs acteurs, dont la DEPP fait partie. Il montre la façon dont l'évaluation, et notamment l'évaluation statistique, est progressivement devenue en un demi-siècle un outil incontournable du pilotage de l'action éducative, trouvant une place de choix dans les outils du Nouveau Management Public. L'arrivée d'un genre nouveau d'évaluations « benchmark » internationales dans les années 1990 contribue à la réorganisation d'un champ pourtant récemment institué afin de laisser une place au banc des experts aux organisations internationales, et notamment à l'OCDE.

Le deuxième chapitre s'applique à retracer le processus de construction de PISA et avec lui l'irruption de l'OCDE dans le champ de l'évaluation des politiques éducatives avec son évaluation PISA, grâce auquel elle se pose en expert avant-garde à la disposition des pays membres pour contribuer à un meilleur pilotage de l'action éducative. A la fois évaluatrice des résultats des politiques éducatives avec une méthodologie et un objet d'évaluation d'un genre nouveau, conseillère en stratégie et experte technique des méthodes statistiques appliquées aux politiques publiques, elle se crée un positionnement tricéphale inédit qui permet de mieux saisir en miroir le repositionnement des évaluateurs nationaux à l'arrivée de PISA.

La deuxième partie porte sur les évaluateurs institutionnels nationaux que sont les agents de la DEPP et analyse la façon dont l'arrivée de PISA remet en question les dimensions de leur expertise pour l'action publique en éducation, récemment acquises et encore instables, et les oblige à se repositionner dans le champ de l'évaluation des politiques éducatives. Désignés interlocuteurs directs et permanents de l'OCDE pour contribuer à l'élaboration puis à la passation des différents cycles de PISA auprès des élèves français, les agents de la DEPP deviennent opérateurs de l'évaluation internationale et représentants des intérêts de la France auprès de l'organisation. Le chapitre trois établit dans un premier temps les marqueurs d'expertise revendiqués par les agents de la DEPP et par les équipes de l'OCDE travaillant à PISA, dans leur pratique de l'évaluation de la politique éducative française et leur rôle d'aide au pilotage de l'action publique. Il interroge en toile de fond le lien entre évaluation et expertise, pour faire apparaître les traits jugés déterminants par les agents de la DEPP impliqués dans PISA pour légitimer leur rôle d'aide au pilotage de l'action publique en éducation et donner corps à leurs prises de position techniques lors de la conception des cycles PISA. Le chapitre quatre étudie l'appropriation par les agents de la DEPP de leur implication jugée par eux inévitable dans PISA, et les stratégies de reconquête de légitimité qu'ils développent dans ce cadre. Derrière une façade de critique de la méthodologie de PISA, la DEPP travaille à construire ses propres évaluations reprenant, pour la plus ambitieuse d'entre elles, des traits centraux de la méthodologie de PISA. Ces travaux d'un genre nouveau s'accompagnent d'un repositionnement des caractéristiques d'expertise de la direction, qui tente de faire valoir, par des arguments nouveaux, la qualité et la robustesse supérieure de ses évaluations et de ses conclusions auprès des décideurs politiques et des membres de sa communauté épistémique. La communication autour de PISA est repensée pour porter un discours unanime de complémentarité des évaluations nationales et internationales, chacune permettant d'éclairer une partie des problèmes publics auxquels les décideurs politiques cherchent à répondre.

La troisième partie analyse dans quelle mesure la stratégie mise en place par la DEPP a eu l'effet escompté à la fois lors du contrôle permanent de l'action du gouvernement et de l'application des lois exercé par les parlementaires, et lors des grands moments législatifs qui ont jalonné notre période. Le chapitre cinq s'intéresse à la mobilisation des évaluations statistiques de la DEPP et de PISA par les parlementaires et met en évidence la façon dont les différences d'usage témoignent du rôle différent alloué par les parlementaires aux évaluateurs nationaux et à l'OCDE. Le chapitre 6 quant à lui regarde au microscope deux grands *moments réformateurs* faisant intervenir au centre de leur processus le principe d'évaluation de l'action éducative passée et à venir : la réforme menée principalement par François Fillon en 2005 introduisant le concept de socle commun de connaissance dont la prégnance à cette époque dans les débats éducatifs est largement attribuée à l'OCDE ; et la réforme menée en 2013 par Vincent Peillon portant comme objectif principal la réduction des inégalités scolaires et tenant les résultats PISA pour preuve de l'urgence de réformer le système éducatif français. Les usages des travaux de la DEPP et de PISA sont très différents, ce soit au moment de construire l'argumentation justifiant l'existence de la réforme, au moment d'en débattre devant la représentation nationale, ou au moment de préparer l'évaluation à venir des résultats des propositions de réorientation faites. Ils témoignent de la difficulté pour la direction du ministère de valoriser son positionnement national, dans un contexte d'internationalisation des débats éducatifs, et ses corollaires comme la formation, le marché du travail, la croissance économique. Le succès de la stratégie de relégitimation des travaux de la DEPP est d'autant plus compliqué à atteindre que l'intensification des usages des évaluations, qui aurait pu permettre de populariser ses travaux, a pour effet collatéral d'interroger la nature du rôle de la DEPP et notamment de questionner sa proximité avec le pouvoir politique à la tête du ministère dans lequel elle est intégrée – et donc l'indépendance de son expertise.

Partie 1 : La construction de PISA, un paradigme à plusieurs égards dans l’histoire des évaluations des politiques éducatives françaises.

Cette première partie vise à inscrire la naissance de PISA et l’organisation de la participation des évaluateurs français à cette évaluation d’un genre nouveau dans le cadre plus large de l’évaluation des politiques éducatives en France. Elle a pour objectif de souligner les particularités du système d’acteurs des évaluations françaises du champ éducatif, la façon dont les rôles des acteurs de première ligne se sont construits et dont les objectifs des démarches d’évaluation ont été assignés. Elle veut montrer ensuite que la récurrence et les singularités de la structuration de ce champ de l’évaluation en éducation en France seront des facteurs cruciaux pour permettre à l’OCDE de construire pour PISA une place singulière dans le paysage de l’évaluation en éducation.

Il s’agit donc d’abord de retracer le processus au cours duquel l’évaluation des politiques éducatives s’est constituée en passage obligé du processus de l’action publique tardivement en France, sous l’influence des théories de nouvelle gestion publique et la constitution de la réforme scolaire comme exigence politique. Le ministère de l’Education nationale a constitué en son sein, après plusieurs batailles idéologiques internes sur la nature des évaluateurs, une direction dédiée à l’évaluation, qui est perçue avant tout comme pourvoyeuse de données statistiques. La DEPP naît à la fin des années 1980, avec pour mission de mettre en place des évaluations statistiques en fonction des orientations politiques décidées par les cabinets ministériels successifs. Si les corps d’inspection deviennent, entre autres missions, les évaluateurs des actions pédagogiques des enseignants, les agents de la DEPP évaluent sur commande du Ministre à toutes les étapes de processus de prise de décision (évaluations d’expérimentations, études d’impact, évaluations cycliques des résultats des élèves, représentation française auprès des organisations en charge d’évaluations internationales, etc.). (Chapitre 1).

C'est dans ce cadre récemment institué que PISA apparaît à la fin des années 1990, développée par une organisation internationale économique à la légitimité forte dans son champ de compétence, mais pas encore dans le champ de l'éducation. En liant économie, développement et éducation, en définissant de façon nouvelle les objectifs de l'éducation et en portant sa vision auprès de ses pays membres, en introduisant le concept de performance et de compétence plus que d'acquisition des connaissances dans sa façon de concevoir l'objet de son évaluation PISA, l'OCDE prend une position originale dans le champ du pouvoir scolaire et dans le champ de l'évaluation de la politique éducative française. Elle se pose en « œil du persan », en évaluation externe complémentaire des études nationales, tout en faisant intervenir à toutes ses étapes de construction l'évaluateur national qu'est la DEPP (Chapitre 2).

Chapitre 1 : L'évaluation en éducation : la naissance tardive d'un besoin de gestion publique en France

« Le pilotage du système éducatif : l'évaluation globale du système éducatif.

Le ministère de l'éducation nationale mène une politique cohérente d'évaluation au niveau national, seul ou avec des organismes spécialisés. L'objectif est d'insuffler, au sein du système éducatif, une culture de l'évaluation fondée sur la recherche de la performance. »

C'est par ces quelques mots liminaires que le gouvernement définit publiquement⁵⁴, en 2019, les missions de pilotage de son action éducative qu'il s'est assigné. Il y intègre les grands concepts au cœur des modèles évaluatifs en vogue dans le secteur privé, et notamment ceux de cohérence et de performance, qui sont ici appliqués à l'ensemble des acteurs du système éducatif. L'évaluation de l'atteinte de ces objectifs par les différents acteurs déterminerait alors la performance du système éducatif dans son ensemble.

Cette définition des modalités de pilotage du système éducatif ne va pas de soi et n'est pas intrinsèquement liée à l'existence d'un « système éducatif français ». Elle correspond à une vision construite par ses acteurs, au premier rang desquels les acteurs politiques, des objectifs qu'ils allouent à l'action éducative⁵⁵. Pour cette raison, au gré des changements politiques, sur fond de bouleversements économiques et sociétaux, l'évaluation de l'éducation a pris bien des objets et des formes depuis la constitution de l'école républicaine à la fin du XIX^e siècle et l'inscription dans le marbre de la loi de ses grands principes directeurs, gratuité, obligation et laïcité. Evaluation des élèves par leurs

⁵⁴ Sur le site gouvernemental education.gouv.fr, consulté le 12 septembre 2017.

⁵⁵ Que nous comprendrons dans l'ensemble de cette thèse comme « *l'action publique en matière d'éducation, et non pas les actions pédagogiques effectuées en classe par un enseignant ou en situation par un éducateur* », selon la définition donnée par Xavier Pons, *Evaluer l'action éducative, des professionnels en concurrence*, Education et société, PUF, Paris, 2010, p.3.

enseignants, mise en place d'évaluations standardisées pour sanctionner la fin d'un cycle scolaire ou permettre l'accès au niveau supérieur, inspection des enseignants, évaluation budgétaire et organisationnelle à tous les niveaux administratifs, le principe de l'évaluation de l'éducation a évolué concomitamment du système éducatif français, sans en présenter dès le départ toutes les dimensions. Ces dimensions plurielles et parfois simultanées de l'évaluation ont été le fait d'une multiplication d'acteurs, dont les intérêts, les compétences et les parcours de chacun ont modelé profondément la façon d'évaluer.

Malgré ses mille visages, l'évaluation n'a été que tardivement institutionnalisée et structurée, dans les années 1980 – près d'un siècle après les lois Ferry –, en un champ d'acteurs évaluateurs aux positionnements et aux rôles définis, mouvant au gré des décisions politiques concernant la place de l'évaluation dans l'action en éducation, et de l'émergence de nouveaux acteurs de l'évaluation. Si l'évaluation était l'outil privilégié des enseignants pour suivre la scolarité de leurs élèves et permettait aux gouvernements de saisir, au niveau national, les résultats globaux de chaque génération d'enfants sortant de l'école, son usage comme outil de pilotage des politiques publiques n'a finalement été institué que très progressivement dans les années d'après-seconde guerre mondiale. Ce surgissement d'usage est un sujet de curiosité : quel fût le contexte propice à son émergence ? Pourquoi, presque soudainement, des années après certains pays partenaires proches de la France, celle-ci s'est-elle emparée de cet outil de pilotage ? Cela a-t-il été à l'origine de bouleversements administratifs, organisationnels, de l'appareil d'Etat français ou est-ce la réciproque ? En d'autres termes, la structure induit-elle la logique d'action ? Ou la logique d'action implique-t-elle des réajustements voire des changements de structures ?

Ces questionnements sont d'autant plus centraux que l'évaluation des politiques éducatives plus spécifiquement est un véritable instrument d'action publique. Loin de n'être qu'un outil de gestion rationnel permettant d'atteindre des objectifs préalablement déterminés, il traduit une représentation propre des problèmes, des rapports de pouvoir entre les acteurs, et permettent de transmettre

des modèles cognitifs⁵⁶. Les instruments sont alors une porte d'entrée de l'analyse de l'action publique, qui sont « *porteurs de valeurs nourries d'une interprétation du social et de conceptions précises du mode de régulation envisagé. (...)Un instrument d'action publique constitue un dispositif à la fois technique et social qui organise des rapports sociaux spécifiques entre la puissance publique et ses destinataires, en fonction des représentations et des significations dont il est porteur* »⁵⁷. Les instruments, par la représentation des problèmes qu'ils portent en eux, donnent un cadre à l'action publique et déterminent les relations entre les acteurs de ce cadre. La déconstruction de cet instrument par le prisme de ses acteurs permet d'en saisir les principes axiologiques.

Pour s'inscrire dans cette démarche, il faut reconstituer un processus analytique à plusieurs niveaux. Définir d'abord des objectifs alloués à l'action éducative, ce qui implique, pour en déterminer la méthode d'évaluation, de définir les objectifs que l'on assigne à l'éducation. L'étude des formes et des acteurs de l'évaluation des politiques éducatives, au-delà de donner seulement des informations concrètes sur l'organisation de l'administration publique, permet de saisir pleinement les enjeux politiques à l'œuvre tout au long de la période. Parmi eux, les jeux d'influence des décideurs politiques bien sûr mais aussi ceux des acteurs de l'évaluation qui, face à l'apparition constante de nouveaux comités, services, directions, voire d'évaluateurs indépendants en dehors du champ administratif, doivent sans cesse se repositionner pour conserver leurs prérogatives et leur visibilité, jouer de leurs outils d'évaluation pour justifier la pertinence de leurs conclusions et s'assurer de recevoir l'écoute attentive des décideurs. L'évaluateur-expert est ici un objet d'étude aux multiples facettes, permettant d'éclairer non seulement la constitution du champ de l'expertise de l'évaluation de l'action éducative, mais qui fait aussi apparaître en creux les sinuosités des enjeux assignés aux politiques éducatives par ceux qui les font et les défont.

⁵⁶ Salvatore Maugeri (dir.), *Délit de gestion*, La Dispute, Paris, 2001, cité dans Pierre Lascombes et Patrick Le Galès, *Gouverner par les instruments*, Paris, Presses de Sciences Po, 2005, p.14.

⁵⁷ Pierre Lascombes et Patrick Le Galès, *Gouverner par les instruments*, Paris, Presses de Sciences Po, 2005, p. 13.

Dans un premier temps nous retracerons donc ici les différentes étapes qui ont construit la tradition d'évaluation de l'action éducative en France, en mettant en lumière les traits qui pourront expliquer ensuite les incompatibilités de conception de l'évaluation (ses objets, ses méthodes, ses cibles, ses usages) entre les évaluateurs institutionnels français et les évaluateurs internationaux de l'OCDE. Dans un second temps, nous analyserons la façon dont l'histoire de la création de la direction en charge de l'évaluation statistique de l'action éducative, interne au ministère de l'Education Nationale, donne à voir le positionnement qu'elle occupe dans le groupe fourni des évaluateurs des politiques éducatives françaises. Il s'agira de voir la façon dont ses missions ont été officiellement définies et dont elles ont en pratique varié selon les personnes à la tête à la fois du ministère et de la DEPP.

1) Une histoire longue de l'évaluation de l'action éducative

Le commencement de l'histoire de l'évaluation de l'action éducative en France est souvent daté du début des années 1980. Près de cent ans après les lois Ferry instituant l'école républicaine en 1881 et 1882, après n'avoir porté que sur le recueil des résultats des élèves aux évaluations nationales sanctionnant la fin des différents cycles d'enseignement, l'Education nationale se serait donné les outils permettant d'évaluer sa propre action. Pourtant l'histoire des évaluations de l'action éducative est en fait plus longue. Les premières d'entre elles remontent au début du siècle et ont posé les jalons de la tradition évaluative française encore appliquée cent ans plus tard. Cette histoire longue de l'évaluation de l'action éducative permet d'éclairer non seulement la période aveugle qui caractérise le pilotage des politiques éducatives française les soixante-dix premières années du XX^e siècle, mais aussi d'expliquer l'origine des méthodologies privilégiées par les évaluateurs français une fois instituées dans les années 1980 ainsi que les rigidités techniques de la France envers les évaluations internationales lancées à la fin du

XX^e siècle aux méthodologies réfutées depuis longtemps par les évaluateurs français.

a) L'évaluation de l'éducation à la croisée des chemins entre psychométrie et psychanalyse

Les missions assignées à l'Education nationale au fil deux derniers siècles ont guidé la formalisation des principes assignés à l'évaluation de l'action éducative. Si elles ont fait l'objet de débats ardents après la Révolution, opposant volontiers instruction publique et éducation nationale, elles se sont finalement posées en consensus avec la création du ministère de l'Education nationale au début des années 1930.

Dès la création du ministère de l'Instruction publique en 1824, de grands objectifs issus des débats révolutionnaires lui sont attribués, sans pourtant faire clairement l'unanimité. Sans être le miroir de l'instruction privée religieuse, l'instruction publique est alors celle qui sort du domestique. Elle est alors comprise au sens d'un cadre d'action dans lequel s'inscrit le programme d'éducation nationale⁵⁸. Pour d'autres les deux notions sont parfois synonymes l'une de l'autre, en particulier dans les discours politiques révolutionnaires⁵⁹. Pour d'autres encore, elles sont accolées l'une à l'autre, enrichissant chacune de ses spécificités l'idéal républicain de l'Ecole d'Etat construisant ses futurs citoyens. Le Député à la Convention nationale Jean Paul Rabaut Saint-Etienne en résume ainsi en 1792 les définitions : « *Il faut distinguer l'instruction publique de l'éducation nationale. L'instruction publique éclaire et exerce l'esprit, l'éducation nationale doit former le cœur.* »⁶⁰ Finalement, l'instruction publique l'emporte sur l'éducation nationale dans les textes, cette dernière étant jugée marquée par un principe de contrainte omniprésente à l'image de celle prévalant dans les camps de travail

⁵⁸ Pierre-Eugène Muller, « De l'instruction publique à l'éducation nationale », dans *Mots*, n°61, décembre 1999. *L'École en débats*, sous la direction de Marie-Anne Paveau et Pierre-Eugène Muller, p.155.

⁵⁹ Dominique Julia, *Les trois couleurs du tableau noir*, Paris, Belin, 1981, p.43.

⁶⁰ Pierre-Eugène Muller, *art cit*, p. 152.

d'éducation commune⁶¹. L'article 22 de la Déclaration des droits de l'Homme et du Citoyen de 1793 énoncera ainsi que « *l'instruction est le besoin de tous. La société doit favoriser de tout son pouvoir les progrès de la raison publique, et mettre l'instruction à la portée de tous les citoyens* ».

Il faut attendre 1932 pour que l'Instruction publique se transforme en ministère de l'Education nationale, une transformation bien loin de se réduire à l'abandon d'une dénomination obsolète. Au contraire, le gouvernement se rappelle à la mémoire des principes de 1793 et des physiocrates du XVIII^e siècle. Le nouvellement nommé ministre de l'Education nationale, Anatole de Monzie, exprime ainsi lors d'un discours son programme pour l'école dans lequel il résume les principes qui constituent pour lui l'Education nationale : « *le principe de l'égalité scolaire, la règle intégrale de gratuité, l'obligation pour l'Etat de trier et d'orienter les aptitudes enfantines et juvéniles*⁶² ». En instituant ainsi la responsabilité de la nation d'offrir à la fois aux enfants une école unique tout en les orientant vers les cursus les plus adaptés, le discours de De Monzie reprend en filigrane un objet de travail du gouvernement sujet à bien des controverses à la création de l'école républicaine trois décennies plus tôt : l'évaluation des aptitudes des élèves afin de les faire accéder aux filières les plus adaptées, lancée par le ministre Bourgeois.

L'histoire de cette évaluation, inédite dans l'histoire de l'action éducative, est intéressante à plusieurs égards. Au-delà des véritables objectifs de fond du changement d'organisation de l'école qu'elle doit préparer, elle est d'abord l'une des premières utilisations de l'évaluation *ex ante* en éducation, commandée dans ce cas par le ministre pour servir la prise de décision publique. Ensuite, elle pose la première pierre de l'utilisation d'évaluations scientifiques des résultats des élèves pour piloter la politique éducative française. La démarche évaluative est d'autant plus intéressante qu'elle permet de comprendre, cent ans plus tard, les premières oppositions des évaluateurs français aux évaluations internationales

⁶¹ Abbé Grégoire, *Discours sur l'éducation commune*, 30 juillet 1793, cité par Dominique Julia, *op. cit.*

⁶² Cité par Pierre-Eugène Muller, *art. cit.*, p.155.

standardisées auxquelles la France prendra part. Elle permet de mieux saisir les racines des oppositions de fond qui ont pu naître entre les évaluateurs et décideurs politiques français et américains, notamment au moment de la création des évaluations PISA et de la définition de la méthodologie d'évaluation choisie pour ce programme dérivée en grande partie des principes psychométriques décrits par les évaluateurs français. Pourtant, en remontant aux débuts de l'école républicaine et aux premières évaluations commandées par le ministère de l'Éducation, alors Instruction publique, on s'aperçoit que les courants psychométriques trouvent leur source dans les changements législatifs de la fin du XIX^e siècle qui voit naître l'école républicaine gratuite, laïque et obligatoire.

Dès la conceptualisation de l'école républicaine et sa mise en application, les premières volontés d'évaluer sont formalisées par le Ministre de l'Instruction publique de l'époque, Léon Bourgeois, à travers la commande d'une mission au psychologue Alfred Binet, dont le travail porte davantage sur l'étude des différences individuelles que sur les thématiques en vogue jusqu'ici au sein de la psychologie française et notamment la psychologie des états de conscience⁶³. Celui-ci est chargé dès 1904 de travailler à l'élaboration d'une échelle de mesure des résultats des élèves et des causes des résultats les moins bons. Cette mission avait été commandée dans ce qui correspondra plus tard, une fois l'évaluation entrée dans les pratiques de gestion publique, à un objectif de pilotage des politiques publiques. Il s'agissait en effet d'étudier les moyens de mise en œuvre de l'obligation d'instruction de tous les enfants⁶⁴ et de préparer la création par la loi de classes de perfectionnement pour les élèves ayant un retard scolaire de plusieurs années par rapport au cursus classique⁶⁵. Binet sera au cœur de la

⁶³ Georges Cognet, « Les trois vies du Binet-Simon : Binet-Simon, Nemi, Nemi-2 », *Recherches et éducatives*, octobre 2011.

⁶⁴ Principe énoncé dans la loi du 28 mars 1882 portant sur l'obligation scolaire pour tous les enfants de 6 à 13 ans.

⁶⁵ Cette réflexion sur la scolarisation des enfants à besoins particuliers a été engagée sous la pression du médecin Désiré Magloire Bourneville, responsable de la Section des enfants de Bicêtre. Il souhaitait, face à la surcharge de son service, externaliser d'une part la prise en charge des enfants « arriérés » en évitant autant que possible de les hospitaliser sauf en cas de pathologies lourdes, leur permettant ainsi de rester proche de leurs familles et de pouvoir suivre une scolarité adaptée. Proche du pouvoir et des inspecteurs d'arrondissements de Paris, Bourneville réussit à faire entendre son projet d'enseignement spécialisé et propose de créer une

construction de l'action éducative. Missionné en amont des décisions législatives par la Commission Bourgeois, il est ensuite chargé d'ouvrir une classe expérimentale, en 1907, après le dépôt du projet de loi proposant la création des classes de perfectionnement.

Les recherches et expérimentations de Binet aboutissent finalement à donner au ministre Bourgeois les outils pour lancer sa réforme. Les classes de perfectionnement sont créées par la loi du 15 avril 1909 et sont placées sous le contrôle de l'Instruction publique. Les instituteurs qui y enseignent sont dotés d'un nouveau statut nécessitant de suivre des formations particulières leur permettant d'acquérir un « Certificat d'aptitude à l'enseignement des enfants arriérés »⁶⁶ (CAEA). Si la création des classes de perfectionnement a longtemps été considéré comme ayant été le fruit d'un raisonnement construit selon lequel l'obligation de scolarisation de tous les enfants aurait engendré un afflux d'élèves en difficulté dans les classes, sans que les enseignants n'y soient préparés, il s'agit en fait davantage d'arguments issus du discours politique justifiant la réforme que du processus de construction effectif de ce nouveau dispositif⁶⁷. La demande n'émanait pas du corps enseignant attaché à l'égalité des enfants devant l'éducation et opposé aux démarches de différenciation des élèves⁶⁸, et la question de savoir comment comprendre les difficultés rencontrées par certains enfants n'était, à l'époque, pas abordée dans le cadre scolaire, restant le fait principalement des psychologues et des médecins.

Alfred Binet, associé au psychiatre Théodore Simon, est missionné par la Commission Bourgeois pour élaborer une méthodologie permettant de trier les élèves selon leurs capacités intellectuelles, en distinguant ceux capables de suivre une scolarité en établissement classique et ceux nécessitant un accompagnement

commission en charge d'évaluer son bienfondé et comprenant des membres de la Direction de l'enseignement. Voir Jacques Arveiller, « La loi de 1909 et la définition de l'instituteur spécialisé, *Les Sciences de l'éducation - Pour l'Ère nouvelle*, n°1 Vol. 42, 2009.

⁶⁶ Le terme « arriéré » était utilisé à l'époque. Il n'est plus employé aujourd'hui et a été remplacé en médecine clinique par « déficient intellectuel léger ».

⁶⁷ Jacques Arveiller, « La loi de 1909 et la définition de l'instituteur spécialisé, *Les Sciences de l'éducation - Pour l'Ère nouvelle*, n°1 Vol. 42, 2009.

⁶⁸ Monique Vial, *Les débuts de l'enseignement spécial en France*, Travaux du CRESAS, n°18, 1979.

particulier⁶⁹. L'objectif est d'identifier, parmi les élèves aux résultats scolaires les plus faibles, ceux dont le retard est lié à un manque de travail et les autres. Binet et Simon définissent une échelle métrique de l'intelligence fondée sur des critères à la fois individuels (un handicap, ce qu'ils appellent la paresse, etc.) et sur des causes extrinsèques (la précarité de la famille, des maladies qui peuvent toucher des enfants). Binet ajoute des causes mixtes, qui rendent plus complexes le repérage des enfants ayant besoin d'un accompagnement spécial.

Cette échelle métrique ne rencontre pas un franc succès en France, si ce n'est ponctuellement pour justifier la mise en place des classes de perfectionnement. Elle est contestée sur le plan des méthodes qu'elle emploie, réfutées par les puissants courants de psychologie scientifique et de psychiatrie de l'époque : mesurer l'intelligence est particulièrement subtil, et les méthodes employées par Binet seraient biaisées par la suggestibilité des enfants, parce qu'elles sont trop proches des exercices scolaires et surtout inadaptées pour pouvoir construire un cadre d'action répondant aux besoins des enfants⁷⁰. La question du score unique interroge, et nombreux sont ceux à s'opposer à cette conception unidimensionnelle de l'intelligence. Par exemple, le courant initié par Henri Piéron, tout en reconnaissant l'intérêt des travaux de Binet, s'attache à construire une mesure pluridimensionnelle de l'intelligence pour l'appréhender en termes de profils plus que de score. Les travaux de Binet et Simon sont rapidement catégorisés comme aidant à l'orientation des élèves plus qu'à l'évaluation des politiques éducatives en amont ou en aval et trouvera une filiation dans le développement du métier de conseiller d'orientation-psychologue⁷¹. Ils connaîtront en revanche un grand succès aux Etats-Unis où ils

⁶⁹ Binet indique ainsi en 1905 : « Notre but est, lorsqu'un enfant sera mis en notre présence, de faire la mesure de ses capacités intellectuelles, afin de savoir s'il est normal ou si c'est un arriéré. », Binet Alfred et Simon Théodore, « Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux », dans *L'année psychologique*, vol. 11, 1904.

⁷⁰ Michel Huteau, « Alfred Binet et la psychologie de l'intelligence », *Le journal des psychologues*, 1, n°234, 2006.

⁷¹ Selon l'un de nos enquêtés chef de service à la DEPP, issu lui-même d'une formation de conseiller d'orientation psychologue, les théories de Binet sont enseignées abondamment lors des formations de COP dans une perspective historique. Il est considéré comme « le père de l'orientation scolaire ».

serviront de base au développement de grandes enquêtes auprès des écoliers dès les années 1910. Le test fait l'objet d'adaptations dans les décennies qui suivent, pour aboutir finalement à une évaluation statistique sophistiquée du quotient intellectuel administré à l'ensemble des élèves américains. Après la Seconde Guerre mondiale, les examens d'entrée aux universités prennent la forme de tests d'aptitudes fondés méthodologiquement sur les mêmes principes que les tests de QI, présentés sous la forme de questionnaires à choix multiples.

On le voit, la tradition française de l'évaluation en éducation n'a pas adhéré aux principes de la psychométrie, contrairement au modèle américain qui a paradoxalement fondé sa conception méthodologique de l'évaluation des acquis des élèves sur la psychométrie « à la française ». A la suite de ce paradigme raté, l'évaluation de l'éducation reste éloignée de la sphère de l'action éducative pendant plusieurs décennies, ne prenant principalement la forme que d'évaluations individuelles des élèves administrées en classe par les enseignants, ou au niveau national dans le cadre des diplômes de fin d'étude.

b) Exigence de réforme scolaire et nouveau management public : La naissance de l'évaluation commandée

Si des prémices d'évaluation utilisées pour penser des politiques éducatives peuvent être observées dès les premières années du XX^e siècle en France, il faudra attendre les années 1970 pour que les premières évaluations de l'action éducative soient menées par le ministère de l'Education nationale. Quant à son institutionnalisation, elle date effectivement du milieu des années 1980 avec la création officialisée en 1987 d'une direction dédiée à l'évaluation au sein du ministère de l'Education nationale, la Direction de l'Evaluation et de la Prospective. Nous comprendrons l'institutionnalisation comme ce qui résulte *« de processus qu'on ne saurait réduire à la codification des dispositifs et des règles, à la production de textes législatifs et de règles, à l'adoption d'un nouveau vocable... Elle officialise et donne en modèle des pratiques et des savoir-faire d'échanges entre acteurs*

(...). »⁷² Prise en ce sens, l'institutionnalisation de la DEP, devenue par la suite DEPP, constitue le point de départ de la définition du rôle de l'évaluation et de sa mission d'accompagnement constant de la prise de décision d'action éducative.

Cette utilisation de l'évaluation comme outil de pilotage des politiques éducatives est relativement tardive en France, notamment en comparaison des pays anglo-saxons et compte tenu de l'ancienneté de l'institution de son système éducatif et de son ouverture à l'ensemble des élèves dès la fin du XIX^e siècle. Si l'on compare son histoire avec celle d'autres pays occidentaux comme les pays nordiques, régulièrement portés au pinacle pour leur capacité d'adaptation et d'action forte pour entreprendre des réformes éducatives jugées indispensables, la France s'est pourtant inscrite dans la même temporalité que ses voisins. La Finlande n'a intégré le rôle fondamental de l'évaluation en éducation dans la loi qu'en 1998, en mettant en place un système évaluatif pour passer du pilotage par les normes à l'objectif de résultats⁷³. De la même façon, l'Allemagne, où les Länder ont depuis l'après-guerre renforcé leur compétence en matière d'éducation, n'intègre l'évaluation dans le champ de vision politique et administratif que dans les années 1990⁷⁴. Ces proximités dans la temporalité des processus d'institutionnalisation de l'évaluation en éducation amènent à deux constats. D'une part, les pays européens, quelle que soit la structure de leur système éducatif et l'ancienneté de la massification de l'accès à l'enseignement, ont initié l'institutionnalisation de l'évaluation de l'éducation à la même époque, répondant à une exigence commune de rationalisation de l'action publique nécessitant de déployer de nouveaux outils d'administration⁷⁵. D'autre part,

⁷² Jacques Lagroye, Bastien François, Frédéric Sawicki, *Sociologie politique*, Paris, Presses de Sciences-Po – Dalloz, 2002.

⁷³ Hannu Simola, Janne Varjo, Risto Rinne, « À contre-courant : dépendance au sentier, convergence et contingence. Vers une meilleure compréhension du modèle finlandais d'assurance-qualité et d'évaluation », *Education et sociétés*, 2, n°28, 2011.

⁷⁴ Brigitte Steinert, « Evaluation et système-qualité, Panorama en Allemagne », *Revue internationale d'éducation de Sèvres*, juin 2000.

⁷⁵ Voir par exemple l'article retraçant l'histoire de l'évaluation des politiques publiques : Patrice Duran, « L'évaluation des politiques publiques : une irrésistible obligation. Introduction générale », *Revue française des affaires sociales*, n°1-2, 2010.

l'inspiration commune de ces nouvelles pratiques est venue de la pratique des pays anglo-saxons de l'évaluation de l'action publique⁷⁶. Ils ont été le moteur de l'introduction de la logique évaluative en éducation, et se sont emparés des principes de la Nouvelle Gestion Publique pour les appliquer au champ éducatif presque dix ans plus tôt. L'exemple américain notamment est d'autant plus intéressant à regarder que les principes de l'évaluation en éducation qui sont appliqués sont directement issus du courant lancé par Binet en début du siècle et qui n'a pas trouvé d'écho en France et en Europe.

i. L'*accountability* américaine

Les Etats-Unis ont été les fers de lance de ce *New Public Management* (NPM), appliqué pour répondre à leur vision de l'intérêt général⁷⁷. Si, comme en France, les fonctionnaires ont pour mission au-dessus de toute autre d'œuvrer pour assurer le bien commun, le statut de leur fonction est bien différent et donne les bases de la construction d'une logique évaluatrice fondée sur les résultats, et non pas sur les moyens employés pour y parvenir. Les civil servants américains n'ont en effet pas de statut spécifique de principe par rapport aux emplois du secteur privé. Leur position est révocable et leur renouvellement est encouragé. Au contraire du système français où la continuité du service public est un principe fondamental de l'administration, et où le statut des fonctionnaires est construit en conséquence : un emploi par principe particulier dont le maintien dans le temps est un critère indispensable pour profiter de l'expérience grandissante des agents, affectés au cours de leur carrière là où l'intérêt général l'exige⁷⁸. Si les

⁷⁶ Jean-Claude Barbier, « Eléments pour une sociologie de l'évaluation des politiques publiques en France », *Revue française des affaires sociales*, n°1-2, 2010.

⁷⁷ Fabien Bottini, « Le new public management : facteur d'un alignement des fonctions publiques américaine et française sur le droit commun du travail ? », *Les politiques du marché du travail face à la crise économique*, 2014.

⁷⁸ Le statut particulier de la fonction publique a été inscrit dans la loi française en 1934 puis réaffirmé dans la loi du 19 octobre 1946 faisant la distinction entre l'emploi et le grade, qui est lié au fonctionnaire et dont il ne peut être séparé que par une procédure lourde applicable dans quelques cas graves. Entre 1984 et 1986, trois textes de loi définissent plus précisément le statut spécifique des trois fonctions publiques, d'Etat, territoriale et hospitalière.

fonctionnaires américains doivent être recrutés uniquement sur un critère de compétence et d'efficacité⁷⁹, quel que soit leur position hiérarchique, l'avancement des fonctionnaires français se fonde principalement, depuis la loi du 27 février 1912, sur leur ancienneté ou leur mérite.

L'exigence d'efficacité de l'action publique et d'atteinte des objectifs fixés s'est d'abord portée sur la vérification et le contrôle, voire la maîtrise, des dépenses publiques dès les premières décennies du XX^e siècle, dans les grands champs d'action de l'Etat « gardien de nuit » : l'armée et le contrôle de la création monétaire⁸⁰. Le *Government accountability office* est créé en juillet 1921, au premier chef afin de contrôler les dépenses publiques jugées confuses depuis la fin de la Première guerre mondiale. Doté de responsabilités grandissantes de décennies en décennies, ce GAO voit ses missions officiellement élargies dès le début des années 1970 et prend alors part à la construction des budgets. Il commence à faire entrer des experts thématiques (issus du monde scientifique ou universitaire en particulier) en son sein pour travailler sur des sujets d'action publique spécifiques et en dehors de la politique économique, de la santé publique aux politiques sociales en général⁸¹.

Dans le champ de l'action éducative, les Etats-Unis mettent en place, dès les années 1980, les premiers tests d'évaluation standardisée à grande échelle. L'idée est alors de produire des outils de pilotage des politiques éducatives, à la fois pour améliorer le fonctionnement du système et les résultats des élèves. Les tests sont fédéraux, dépassant la marge de manœuvre individuelle des Etats pourtant très attachés à leurs *Department of Education* qui définissent l'ensemble de la politique scolaire menée sur leur territoire, des programmes aux habilitations des enseignants⁸². Les Etats deviennent alors les « administrateurs

⁷⁹ Un principe inscrit dans la loi américaine du 10 janvier 1833, le *Civil Service Act*, ou *Pendleton Act*.

⁸⁰ François-Xavier Merrien, *L'Etat-providence*, Paris, PUF coll « Que sais-je », 2007 [3^e éd.], p.3-4.

⁸¹ Il faudra par exemple attendre la réforme constitutionnelle de 2008 en France pour que l'homologue du GAO, la Cour des Comptes, se voit accorder un rôle d'accompagnement du Parlement dans le contrôle de l'action du gouvernement.

⁸² Laurence Emile-Besse, « Le système éducatif américain », *Revue internationale d'éducation de Sèvres*, n°35, avril 2004.

de test » définis au niveau fédéral et identiques sur tout le territoire. Des objectifs évaluatifs d'efficacité des moyens employés (humains, financiers, matériels) et d'équité entre les élèves sont alors associés à ces évaluations qui étaient jusque-là pensées uniquement en termes de mesure des compétences individuelles des élèves⁸³. Au-delà de la méthodologie psychométrique employée dans les tests qui, par la place qu'elle accorde aux statistiques, est en elle-même une singularité du système anglo-saxon, l'évaluation de l'éducation américaine se dote d'objectifs issus du monde managérial. Cette démarche nouvelle sera par la suite progressivement adoptée par les acteurs de l'action publique française, dans un contexte favorable à son adaptation au fonctionnement politique et administratif national.

ii. Du changement administratif à la réforme de l'Etat : un nouvel impératif de performance pour l'action publique française

En France, l'irruption à la fois de la statistique et des objectifs du Nouveau Management Public sont plus tardifs. Ils ont répondu à des impératifs nouveaux pour l'action publique française directement liés aux évolutions sociales et économiques de la période d'après-guerre. Ces changements sociétaux ont induit des réformes éducatives structurelles nombreuses, qui se sont combinées à des exigences inédites de maîtrise budgétaire. Finalement l'administration et le politique ont vu en les principes du Nouveau Management Public un moyen d'engager les réformes qu'ils considéraient nécessaires, en conciliant dans le même temps transparence démocratique et optimisation de la dépense publique en éducation.

Les années 1970 ont plus particulièrement vu se développer des réformes de l'administration. La littérature à leur sujet identifie plusieurs facteurs qui, une

⁸³ Cécile de Bouttemont, « Etats-Unis : les tests d'évaluation en débat », *Revue internationale d'éducation de Sèvres*, n°53, avril 2010.

fois combinés, les ont rendu nécessaires.⁸⁴ Au premier rang d'entre eux, la crise budgétaire, fiscale et financière s'accompagne ou dérive du poids toujours plus grand de l'Etat notamment en matière de dépenses publiques. Les crises des années 1970 engendrent l'apparition d'un problème nouveau, ou plutôt d'un phénomène nouvellement problématique : la croissance du déficit public. La massification de l'accès au système éducatif public, gratuit, entraîne mécaniquement une hausse des dépenses de l'Etat. La dépense publique annuelle d'éducation a ainsi augmenté six fois plus vite que le PIB au XX^e siècle. En 1980, les dépenses intérieures d'éducation (DIE), qui regroupent selon la définition donnée par l'INSEE « *toutes les dépenses effectuées par l'ensemble des agents économiques pour les activités d'éducation* », atteignent près de 30 milliards d'euros dont 68% sont à la charge de l'Etat. Elles augmentent jusqu'au milieu des années 1995 pour atteindre au plus haut 7,6% du PIB en 1997, contre 1% en 1901. En 1989, le budget de l'éducation nationale devance tous les autres : il est le budget le plus important au sein de l'Etat, devant celui des anciennes prérogatives de « gardien de nuit » de l'Etat avant l'Etat-providence – la défense nationale notamment. Les instances de contrôle de la dépense publique ne s'intéressent qu'à cette période à l'évaluation de la performance de l'action publique. Il faut en effet notamment attendre la fin des années 1980 pour que la Cour des Comptes s'empare de sujets éducatifs et publie ensuite quelques rapports sur le fonctionnement du système⁸⁵.

L'administration connaît par ailleurs à cette époque une crise d'efficacité et peine à atteindre ses objectifs. Les défaillances identifiées sont nombreuses : manque de connaissances spécifiques pour pouvoir atteindre les objectifs de façon satisfaisante, redondances et lourdeurs administratives, manque de coordination entre les acteurs, manque de considération pour les administrés...

⁸⁴ Ces facteurs sont largement décrits par Philippe Bezes et en particulier dans son ouvrage référence *Réinventer l'Etat, les réformes de l'administration française (1962-2008)*, Paris, PUF, 2009.

⁸⁵ Par exemple Cour des Comptes, *La gestion des enseignants du premier degré*, 1990 ; Cour des Comptes, *L'organisation du baccalauréat et les IUT*, 1994 ; Cour des Comptes, *Les effets de la décentralisation dans l'enseignement secondaire*, 1995.

Dans le champ de l'éducation, on l'a vu, les différents changements mis en œuvre les deux décennies précédentes pour démocratiser l'enseignement secondaire n'ont pas porté les fruits attendus, sans que des études précises pour en expliquer les raisons soient concrètement menées. Finalement, l'évaluation *ex post* apparaît aux décideurs comme un moyen de saisir les raisons de ces changements en demi-teinte.

De cette crise d'efficacité découle une crise de légitimité, où les nombreux échecs en matière de politique publique entraînent une érosion de la légitimité accordée par les citoyens aux gouvernants. Les décideurs politiques risquent alors d'être perçus comme davantage préoccupés par leur dynamique de carrière et les réélections possibles que par la résolution de problèmes administratifs ou les propositions de politiques publiques pour en améliorer les effets. Cette crise conduit les décideurs politiques à vouloir partager la responsabilité des échecs et les risques avec les acteurs administratifs. Du côté de ces derniers, une plus grande immédiateté des rapports entre les administrés et les administrations se développe, renforcée par le développement de nouvelles technologies et de mutations sociales qui y sont associées, notamment l'apparition du numérique. Il devient alors indispensable de pouvoir répondre aux questions posées dans le cadre du débat démocratique, et l'évaluation constitue un outil de choix pour apporter des arguments fondés aux oppositions et au scepticisme. Les outils numériques seront, nous le verrons ensuite, un facilitateur du développement d'évaluations statistiques poussées et de grande ampleur menées notamment par les évaluateurs du ministère de l'Éducation nationale, puisqu'ils permettront de traiter efficacement et rapidement un volume important de résultats de l'évaluation des élèves.

Dans la période suivante, le contexte économique défavorable des années post-Trente Glorieuses et la crise de confiance en les gouvernements à la même époque se sont aussi accompagnés de réformes structurelles de l'administration dès les premières années de la décennie 1980. En 1981, les lois Defferre renforcent le poids des collectivités territoriales dans la prise de décision publique et rapprochent les citoyens des lieux de décision en accordant aux conseils généraux

et aux Régions une autonomie nouvelle et de nouvelles compétences. La proximité des centres décisionnels, ajoutée à tous les autres facteurs, a contribué à renforcer les exigences des citoyens envers l'action publique⁸⁶.

Ces nombreux facteurs ont été le terreau fertile de l'apparition dans les années 1980 du nouveau management public, dérivé de la pensée néo-libérale, dans les principes d'organisation de l'administration française.

Les principes de la nouvelle gestion publique sont empruntés à plusieurs disciplines au premier rang desquelles le management et la gestion dans le secteur privé. Malgré leurs origines disciplinaires disparates qui auraient pu multiplier leurs objectifs, les principes de Nouveau Management Public agissent et se combinent dans une direction commune, celle de promouvoir un « mieux d'Etat » en introduisant la logique entrepreneuriale au cœur de l'action publique⁸⁷. Pour bien comprendre les enjeux induits par l'irruption du Nouveau Management Public dans la culture de l'administration française, encore faut-il définir ce qu'on entend par logique entrepreneuriale. Si l'on regarde les différents travaux publiés sur le sujet, y compris en sciences de gestion, la définition reste floue. Difficile de déterminer unanimement ce qui relève de l'entrepreneuriat. La notion est souvent entendue, y compris par les institutions internationales comme la Commission européenne, comme l'ensemble dominant de valeurs liées à l'industrie et l'économie (l'identification des opportunités possibles entraînant la création d'activité notamment), reléguant au second plan les valeurs sociales⁸⁸. Cette définition pose question dans le cadre de l'action publique et a fortiori de l'action éducative, chargées d'œuvrer dans la recherche du plus grand bien commun, du plus grand intérêt collectif. On l'a vu en regardant les spécificités de l'administration française par rapport aux pays anglo-saxons et notamment les

⁸⁶Brice Gaillard, *La performance de l'action publique territoriale*, thèse pour le doctorat de science politique sous la direction de Pascal Jan, soutenue à l'Université de Bordeaux le 16 décembre 2016 ; René Demeestere et Gérald Orange, « Gestion publique : qu'est-ce qui a changé depuis 25 ans ? », *Politiques et management public*, vol. 26/3, 2008.

⁸⁷Fabien Bottini, « L'impact du new public management sur la réforme territoriale », *Revue française de droit administratif*, Dalloz, 2015, p. 717.

⁸⁸Catherine Léger-Jarniou, « Développer la culture entrepreneuriale chez les jeunes », *Revue française de gestion*, vol.5, n°185, 2008.

Etats-Unis, la recherche du bien commun s'appréhende en France grâce à une logique de continuité de l'action des agents, l'inscription dans une carrière sécurisante qui leur permettrait de disposer de la stabilité nécessaire pour penser les moyens d'atteindre cet objectif ultime. L'irruption des valeurs issues du monde économique vient bouleverser cette logique d'action en déplaçant les obligations de moyens vers une obligation de résultats, difficile à rendre compatible avec la construction des carrières administratives à la française⁸⁹. Les maîtres-mots deviennent alors ceux auxquels les anglophones se réfèrent comme les trois « E » : *economy*, *efficiency* et *effectiveness*, soit, en français, économie, rendement et efficacité⁹⁰. Autrement dit, le secteur public devrait atteindre un résultat optimal compte tenu des moyens humains et financiers déployés et « créer de la valeur ». Si dans le cas d'une entreprise la création de valeur correspond à l'augmentation année après année du profit réalisé, dans le cas du secteur public elle est plutôt entendue comme la maximisation, à moyens donnés, des services rendus aux citoyens⁹¹. L'action publique doit devenir « performante » pour répondre aux attentes et aux besoins de l'ensemble des usagers. Loin de n'entraîner qu'un recalibrage des pratiques professionnelles de l'administration, l'irruption des méthodes du Nouveau Management Public incite à recourir à l'expertise au détriment du professionnalisme qui était jusqu'alors le modèle dominant au sein de la bureaucratie française⁹².

Dès lors, il devient incontournable de penser et de choisir les outils pour évaluer, selon des indicateurs le permettant, l'adéquation entre les résultats

⁸⁹ Philippe Bezes, Didier Demazière, Thomas Le Bianic, Catherine Paradeise, Romuald Normand, Daniel Benamouzig, Frédéric Pierru et Julia Evetts, « New Public Management et professions dans l'État : au-delà des oppositions, quelles recompositions ? », *Sociologie du travail*, vol.53 – n°3, 2011.

⁹⁰ Fabien Bottini, « Le new public management : facteur d'un alignement des fonctions publiques américaine et française sur le droit commun du travail ? », *Les politiques du marché du travail face à la crise économique*, 2014 ; Charles Polidano, « Administrative reform in core civil services : application and applicability of the new public management » chapitre dans Willy Mc Court et Martin Minogue (ed.), *The Internationalization of Public Management*, chapter 3, Edward Elgar Publishing. 2001.

⁹¹ CREDOC, Performance, efficacité, efficience : les critères d'évaluation des politiques sociales sont-ils pertinents ? Cahier de recherche, n°299, décembre 2012.

⁹² Philippe Bezes, Didier Demazière, Thomas Le Bianic, Catherine Paradeise, Romuald Normand, Daniel Benamouzig, Frédéric Pierru et Julia Evetts, *art cit.*

obtenus et les objectifs, ainsi que le cadre de l'évaluation. Quel(s) public(s) cible(s) ? Quelle périodicité des évaluations ? Quelle ampleur d'évaluation ? Dans le cas du secteur public contrairement à ce qui peut se pratiquer dans le secteur privé, la question de la périodicité des évaluations se pose tout particulièrement. L'impact des décisions de politique publique ne se fait que rarement par des objectifs immédiats, et une évaluation à court terme ne peut que porter sur les réalisations, et non sur les résultats qui seront eux visibles à moyens et long termes. Cette temporalité longue est d'autant plus vraie dans le champ éducatif où il est souvent nécessaire d'attendre qu'une génération d'élèves soit passée pour réellement mesurer – avec toutes les précautions méthodologiques liées à la pluralité des facteurs intervenant dans l'éducation des enfants – les effets d'une politique. La nouvelle gestion publique est donc d'autant plus délicate qu'elle se calque bien souvent sur la temporalité budgétaire d'une année, et nécessite donc le développement d'indicateurs de court terme qui se doivent d'être considérés comme des jalons intermédiaires d'une évaluation plus complète, à même de servir d'outil de pilotage.

iii. La RCB, entre échec de modernisation de l'action publique et institutionnalisation de l'évaluation ministérielle

Malgré cette difficulté structurelle, la construction en un ensemble complémentaire des principes du Nouveau Management Public a facilité leur diffusion d'abord dans les institutions et organisations internationales économiques et financières, du Fonds monétaire international à la Banque mondiale, en passant par l'OCDE. Par leur lobbying auprès de leurs pays membres, ces organisations ont contribué à leur faire adopter l'idée d'adapter ces principes au secteur public. La France tente rapidement de les appliquer à travers la création, en 1968, d'un nouveau mode de planification budgétaire, la rationalisation des choix budgétaires (RCB).

Inspiré du PPBS (*Planning Programing Budgeting System*) en vigueur aux Etats-Unis, la RCB développait non seulement une nouvelle organisation en budgets dédiés à des programmes, mais aussi réfléchissait aux indicateurs pouvant permettre d'évaluer les résultats des actions au sein de chaque programme d'action. L'objectif de ces évaluations était alors bien de fonder les décisions de dépenses à partir de l'identification des problèmes et de la prévision de leurs impacts à la fois sociaux et économiques⁹³. La démarche était pilotée par le ministère des Finances, chaque étude étant délocalisée au sein de son ministère de référence, dans des cellules RCB dédiées. Contextuellement, la RCB s'inscrivait alors dans un mouvement de modernisation des administrations, faisant suite notamment à l'émergence des outils numériques permettant de repenser la gestion des services (appareils informatiques, systèmes d'informations, etc.). Finalement, le lancement de la RCB a abouti à la mise en route d'un demi-millier d'études dont 40 ont été portées en interministériel. Et pour les réaliser, une centaine de chargés d'étude ont été recrutés par les ministères, en interne. Il s'agit là de l'entrée la plus structurante de responsables d'évaluation en interne des ministères, dont la présence survivra à la disparition brutale de la RCB en 1984.

Malgré un développement et un investissement organisationnel d'une telle ampleur, la RCB ne sera qu'une expérimentation longue qui sera arrêtée pratiquement du jour au lendemain. Elle n'aura pas réussi à rendre plus facile la coopération entre les différents ministères, ni à dépasser les difficultés engendrées par les crises économiques des années 1970, obligeant à revoir les budgets en cours d'exercice. Toutefois, deux évolutions intéressantes subsisteront et sont toutes relatives à l'évaluation de l'action publique : d'une part les chargés d'évaluation en interne ne seront pas remerciés malgré l'arrêt de la RCB ; d'autre part l'évaluation ex post apparaîtra comme la solution de pilotage la plus adaptée, au-delà de la seule préparation du budget. Le rapport rendu par Patrick Viveret,

⁹³ Bernard Perret, « De l'échec de la rationalisation des choix budgétaires (RCB) à la loi organique relative aux lois de finances (LOLF) », *Revue française d'administration publique*, vol.1, n°117, 2006.

commandé par le gouvernement de Michel Rocard en 1988, insiste ainsi sur l'apport de l'évaluation au débat démocratique et sur son effet sur le pilotage de l'action publique. Cette volonté d'inscrire l'évaluation dans les pratiques de l'action publique apparaît dans les toutes premières lignes de la lettre de mission adressée à Patrick Viveret :

« Le Gouvernement attache une importance particulière à l'évaluation des politiques publiques. Celle-ci constitue en effet un moyen privilégié pour intégrer la dimension du moyen terme⁹⁴ tant dans l'action de l'Etat que dans le débat et l'intervention des acteurs de la société politique et de la société civile. Elle s'inscrit d'autre part dans un effort global que j'entends conduire pour moderniser les outils d'intelligence politique de notre pays. »⁹⁵

Michel Rocard va jusqu'à inscrire le « *devoir d'évaluation* » dans les quatre orientations principales de ce qu'il nomme la « *politique de renouveau du service publique* »⁹⁶. La « modernité » passe par l'évaluation qui fournit à l'Etat de nouveaux outils de travail à la fois compréhensifs et prospectifs.

Les principes de nouvelle gestion publique ne seront pas abandonnés avec l'arrêt de la RCB, et cette tendance de fond trouvera son apogée avec le lancement quinze ans plus tard d'une transformation totale de la gestion publique française qui nécessitera plus de cinq ans de mise en œuvre : la LOLF.

iv. L'ère des indicateurs de performance : l'instauration de la LOLF

A partir des années 2000, l'efficacité du système et le respect de la contrainte budgétaire sont liés dans la plus pure lignée de la doctrine de nouvelle gestion publique. En reprenant les principes de gestion managériaux du secteur

⁹⁴ Tous les passages soulignés dans cette thèse le sont par moi sauf mention du contraire.

⁹⁵ Lettre de mission du Premier ministre à Patrick Viveret, chargé de mission auprès du Commissaire au Plan, portant commande de propositions en vue de l'évaluation du revenu minimum d'insertion, 22 juillet 1988.

⁹⁶ Dans la Circulaire ministérielle du 23 février 1989 relative au renouveau du service public, JORF, 24 février 1989. Les trois autres orientations sont la mise en place d'une « *politique de relations du travail renouvelée, une politique de développement des responsabilités, une politique d'accueil et de service à l'égard des usagers* ».

privé, les institutions publiques ont peu à peu mis en place des indicateurs de performance accompagnés d'objectifs à atteindre et de nouveaux instruments de mesure pour chacun d'entre eux. Le secteur éducatif n'a pas échappé à ces nouveaux outils et, de la même façon que les autres secteurs de l'action de l'Etat, a été concerné par la réforme majeure des finances publiques mise en œuvre en intégralité à partir de 2006 ⁹⁷: la Loi Organique relative aux Lois de Finances (LOLF). La LOLF, votée le 1^{er} août 2001, présentait trois grands objectifs selon les mots de la direction du budget du ministère des Finances : un objectif de réforme du cadre de la gestion publique pour permettre une orientation vers les résultats obtenus, un objectif de transparence des informations budgétaires et de contrôle parlementaire des comptes publics, et enfin un objectif de facilitation du débat stratégique portant sur les finances publiques.

Le lancement de cette réforme de la conception de la gestion de l'Etat d'une ampleur inégalée s'est faite au grand damne des syndicats qui, à l'époque, l'ont qualifiée d' « *institutionnalisation d'un laminage systématique du service public* »⁹⁸, ne servant qu'à assigner à l'éducation l'unique objectif de baisse des dépenses publiques sans chercher à identifier les besoins des populations. En résumé, la LOLF serait l'outil de transformation du service public en entreprise privée, dans laquelle la seule réduction des coûts à tout prix serait recherchée. Le chef d'établissement deviendrait un gestionnaire, et l'enseignant un salarié. On voit bien ici surgir toute l'ambiguïté de l'application des principes du Nouveau Management Public au secteur public. La logique entrepreneuriale chère à ses penseurs, qui pouvait s'adapter aux objectifs de l'action publique, c'est-à-dire comme on l'a vu la recherche du bien commun, en ne lui demandant « que » de se faire dans des conditions budgétaires optimales, se serait en pratique transformée en « *entreprisation* » du service public dépouillé de son objectif

⁹⁷ La LOLF a été mise en œuvre progressivement à partir du 1^{er} janvier 2002 et a été entièrement appliquée le 1^{er} janvier 2006. Ces cinq années d'entrée en vigueur de la loi témoignent bien de l'ampleur de la tâche qu'ont dû réaliser les différents ministères pour passer d'un budget affecté à chacun à un fonctionnement en missions déclinées elles-mêmes en 150 programmes « concourant à une politique définie » selon l'article 7 de la loi.

⁹⁸ François Weisner, *La LOLF*, document publié par le syndicat SUD éducation le 14 décembre 2004.

intrinsèque. Le chiffre prendrait le pas sur le bien commun, les résultats atteints (et évalués par le biais d'indicateurs statistiques) deviendraient le critère d'évaluation des actions des agents en lieu et place des moyens mis en œuvre.

Pour l'Etat, il s'agit au contraire « *d'instaurer une gestion plus démocratique et plus efficace des dépenses publiques* »⁹⁹, en se donnant des objectifs et en y allouant des moyens, discutés publiquement à chaque fin d'année civile au parlement. L'idée est d'assurer la performance du système en atteignant chaque objectif avec les moyens attribués, voire avec moins. La LOLF est partiellement expérimentée d'abord à partir du 1^{er} janvier 2004 au sein du ministère de l'Education nationale dans deux académies, celle de Bordeaux et de Rennes¹⁰⁰. Les objectifs et les indicateurs correspondant sont définis à partir du mois de juin de la même année. Dix années plus tard, le constat de la Cour des Comptes sur les indicateurs mis en place est sévère. Elle considère que les évaluations fournies pour alimenter la LOLF sont « de faible apport ¹⁰¹ » et ne permettent pas d'avoir une idée précise de l'efficacité du système éducatif. Les indicateurs choisis pour suivre et justifier le budget du ministère seraient trop globaux, ne permettant pas de saisir finement les choix effectués, en particulier financiers. La lecture des enjeux des politiques éducatives est alors faite à travers le prisme de son budget année après année, sans qu'il soit aisé de saisir la construction des raisonnements ayant abouti à la prise de telle ou telle décision de politique éducative. Il est, par ailleurs, intéressant de noter que la DEPP, pourtant organe statistique et évaluateur officiel du ministère, ne fournit que peu de données pour alimenter les indicateurs de la LOLF, en tout cas peu d'indicateurs qualitatifs. Elle s'en tient à l'évolution prévisionnelle des effectifs ou des taux d'accès aux diplômes. La Cour des

⁹⁹ Présentation de la LOLF sur le site <https://www.education.gouv.fr/pid23-cid31/la-lolf-est-que-est.html>.

¹⁰⁰ Seuls les programmes « enseignement scolaire public du premier degré », « enseignement scolaire public du second degré et « soutien de la politique de l'Education nationale », hors emplois et rémunérations des enseignants sont expérimentés. En 2005 la totalité des programmes des premier et second degrés de l'enseignement public sont expérimentés dans les deux académies, et toutes les académies expérimentent le programme « enseignement scolaire public du premier degré ».

¹⁰¹ Cour des Comptes, *L'Education nationale : organiser son évaluation pour améliorer sa performance*, décembre 2017, p. 8.

Comptes regrette que ne puissent être intégrés à la LOLF des indicateurs de plus long terme, et notamment les résultats des évaluations sur les acquis des élèves en fin de scolarité obligatoire, qui ne sont pas renseignés chaque année. Le pilotage budgétaire par l'évaluation se trouve ici face à sa contradiction : une quadrature du cercle impossible à résoudre entre la temporalité des budgets et la temporalité de l'éducation.

Les évaluations de l'action éducative ne trouvent pas une place satisfaisante dans le cadre du contrôle annuel du budget. Elles se sont toutefois positionnées lors des trente dernières années comme un outil apprécié des décideurs politiques pour orienter les réformes, poussées conjointement par le contexte politique, économique et social.

v. Un projet réformateur en prise avec le contexte historique

Le ministère de l'Éducation nationale n'a bien sûr pas échappé à l'irruption des principes du Nouveau Management Public dans l'évaluation de l'action publique, malgré les oppositions syndicales. Il a intégré progressivement les principes de l'évaluation à des fins de pilotage de l'action éducative, portée à la fois par l'évolution de la doctrine de nouvelle gestion publique et par les nouveaux objectifs assignés à l'éducation liés au contexte historique de l'après-guerre.

L'évaluation de l'éducation a répondu à différents enjeux politiques depuis la fin de la Seconde guerre mondiale¹⁰² et à différentes conceptions des effets souhaités des politiques éducatives. La première période, qualifiée par certains auteurs de « *temps de l'égalité* », est celle de l'après-guerre, qui voit progressivement se démocratiser l'accès au système éducatif¹⁰³. La massification de l'école et

¹⁰² Romuald Normand, *Gouverner la réussite scolaire. Une arithmétique politique des inégalités*, Berne : Peter Lang et Lyon : ENS de Lyon, 2011.

¹⁰³ Antoine Prost, *Du changement dans l'école. Les réformes de l'éducation de 1936 à nos jours*, Paris, Seuil, 2013, p. 118 cité par Florence Defresne et Jérôme Krop, « La massification scolaire sous la V^e république, une mise en perspective des statistiques de l'Éducation nationale (1958-2014) », *Éducation et formations*, n°91, septembre 2016.

l'orientation des élèves sont défendues d'abord pour faire face aux besoins en main d'œuvre qualifiée du pays. Si elle avait été initiée dans les textes, la réalité était bien différente. L'accès à l'enseignement secondaire est resté, malgré plusieurs lois censées l'ouvrir aux élèves des milieux modestes (gratuité du collège et du lycée en 1930, suppression des classes élémentaires en lycée accessibles aux enfants des milieux aisés en 1945, etc.), l'apanage des enfants issus de milieux aisés jusqu'aux années 1960. Un chiffre en témoigne : lors de l'année scolaire 1957-1958, 4 millions d'élèves sont scolarisés en primaire élémentaire mais ils sont seulement 70000 dans les premières années de lycée. De nombreuses réformes se succèdent pour ouvrir l'accès au collège. La réforme Berthoin de 1959 généralise l'entrée en sixième et pose les fondements du prolongement de la scolarité obligatoire jusqu'à 16 ans. La loi Haby votée en 1975 et entrée en application deux ans plus tard crée le collège unique et fusionne les filières tout en maintenant un palier d'orientation à la fin de la cinquième pour orienter une partie des élèves vers des CAP en trois ans.

C'est dans ces années 1970 que Claude Seibel, Inspecteur général de l'INSEE, ancien élève de l'école polytechnique et de l'ENSAE¹⁰⁴, responsable des V^e et VI^e Plans en tant que responsable du Service des Programmes de l'INSEE à partir de 1970, crée en 1973 le Service des études informatiques et statistiques du ministère de l'Éducation nationale¹⁰⁵. Au-delà de la prévision de l'évolution des effectifs, il instaure des évaluations des résultats des élèves à partir de panels, à des fins d'étude. Un premier panel de collégiens recrutés à l'entrée en 6^e est mis en place en 1973. Il est suivi en 1978 par un panel d'écoliers recrutés à l'entrée au cours préparatoire. Ces études longitudinales permettaient d'étudier les trajectoires des élèves et leur parcours avec des objectifs très en lien avec les récentes réformes éducatives mises en œuvre. Alors que de multiples changements de fonds du fonctionnement du système éducatif ont été instaurés depuis le début des années 1970, l'objectif de ces études étaient de mieux saisir

¹⁰⁴ Ecole Nationale de la statistique et de l'administration économique de Paris, créée en 1942 sous le nom d'Ecole d'application de la statistique, puis d'Ecole d'application de l'INSEE lorsque l'INSEE est créé en 1946. Elle forme les administrateurs de l'INSEE.

¹⁰⁵ Il créera en 1993 la direction chargée des études du ministère du travail, la DARES.

leur effet sur les cursus des élèves et de donner des outils pour évaluer l'atteinte ou non de l'exigence d'égalité des chances des élèves lors de leur scolarité¹⁰⁶.

Ces premiers panels seront pionniers dans ce qui sera une méthode de choix pour évaluer l'action éducative jusqu'aux années 2000¹⁰⁷. Ils seront reconduits plusieurs fois dans les décennies qui suivront, tous en lien avec des moments de changements du système éducatif. On peut ainsi dire que le panel de 1973 s'est inscrit dans la lignée de la généralisation de l'accès au premier cycle de l'enseignement secondaire ; celui de 1978 au développement de l'accès à l'école maternelle à trois ans ; celui de 1980 (un panel de collégiens) dans la mise en place du collège unique ; celui de 1989 (des collégiens également) dans la mise en œuvre de l'objectif de faire atteindre le baccalauréat à 80% d'une classe d'âge ; celui de 1995 (des collégiens encore) à la réforme du collège de 1994 ; celui de 1997 (des écoliers cette fois) à la mise en place des cycles à l'école élémentaire ; celui de 2007 (des collégiens de nouveau) à la loi d'orientation de 2005 et à la réforme de l'éducation prioritaire de 2006.

Ces panels ne font pas réellement intervenir d'évaluation des acquis des élèves qui les constituent. La « réussite » de la réforme n'est alors pas mesurée par les résultats des élèves qui ont pu en bénéficier mais par leur parcours scolaire. Des remontées de résultats aux évaluations nationales et de notes aux examens sont prises en compte à partir de 1989 mais pas avant. Pourtant quelques évaluations de ce type avaient été menées dès 1975 par le service d'informations économiques et statistiques. Elles portaient alors sur les acquis des élèves en fin d'année de 5^e. Un petit sous échantillon du panel de 1978 avait également été soumis à une évaluation de leurs acquis à l'école primaire, sans que les résultats ne soient intégrés aux données, notamment sociodémographiques, recueillies

¹⁰⁶ Dès le début des années 1960, dans la lignée des réformes Berthoin, deux chercheurs de l'INED, Girard et Bastide, avaient étudié les trajectoires d'élèves après le CM2 grâce à cette méthodologie des panels et avaient montré que les cursus suivis dans le secondaires étaient très divers et fortement marqués par une sélection sociale.

¹⁰⁷ Jean-Paul Caille, « Quarante-cinq ans de panels d'élèves à la Direction de l'évaluation, de la prospective et de la performance (DEPP), *Education et Formations*, Les panels d'élèves de la DEPP : source essentielle pour connaître et évaluer le système éducatif, 2017.

dans le cadre du panel. Les deux évaluations suivent des chemins parallèles et jamais ne se croisent. Il faudra attendre une décennie pour que les acteurs de l'action éducative s'intéressent de plus près aux évaluations des acquis des élèves, qui prendront dès le départ la forme d'évaluations standardisées.

vi. L'avènement des évaluations standardisées

L'évaluation phare devient à la fin des années 1980 l'évaluation standardisée, souvent nommée en interne « test national » ou « évaluation nationale ». Conçue et administrée par une autorité éducative centrale, elle est méthodologiquement similaire dans l'ensemble des pays qui la pratiquent. L'évaluation standardisée évalue à un instant T les résultats d'un ensemble d'élèves d'un même niveau scolaire à un même examen. Elle est passée dans des conditions les plus identiques possibles auprès de tous les élèves concernés afin de réduire les biais de passation. Si les grands principes méthodologiques (représentativité de la base répondante, unicité des tests et passation simultanée) sont les mêmes dans tous les pays qui les pratiquent, les évaluations standardisées ne sont pas conçues selon les mêmes questions évaluatives et n'ont pas toutes le même objectif d'usage.

Trois grandes catégories de tests nationaux peuvent être faites, selon les questions évaluatives et les objectifs de politique publique qu'ils portent. Le premier correspond aux tests qualifiés de « sommatifs » qui sont réalisés à la fin d'un cycle ou d'une année scolaire spécifique. Ils évaluent les apprentissages avec une certification à la clé permettant la comparabilité des diplômes obtenus ou des niveaux scolaires atteints, ou aident à prendre des décisions pédagogiques concernant le parcours scolaire des élèves. Ils ne sont pas utilisés pour évaluer l'action des enseignants mais bien l'acquisition de compétences ou de connaissances par les élèves et n'induisent pas de prise de mesures politiques nationales pour influencer sur la pédagogie en classe. Le second type d'évaluation donne des éléments de réponse pour évaluer le système éducatif dans son ensemble ou les établissements scolaires qui le constituent. Les résultats des tests

sont un indicateur de la performance de ce qui est évalué, de l'atteinte des objectifs assignés, de l'utilisation équitable des moyens alloués. Le troisième groupe d'évaluations, celui des évaluations dites « formatives », a pour objectif d'accompagner l'apprentissage des élèves en identifiant leurs besoins individuels. Finalement, les tests sommatifs et formatifs, qui ont été les premiers à être mis en place en France, n'ont jamais été considérés par le ministère de l'Éducation nationale, ou avant lui le ministère de l'Instruction publique, comme ayant une portée évaluative de l'ensemble du système éducatif, et encore moins de l'action éducative. Ils ont, comme on a pu le voir avec l'exemple des tests Binet-Simon, été employés à des fins d'orientation des élèves dans les filières que leurs enseignants jugeaient adaptées à leurs capacités, ou à évaluer le nombre d'élèves atteignant un niveau scolaire jugé acceptable pour leur permettre d'avoir acquis les outils pour entrer dans la vie d'adulte et de citoyen.

La première évaluation standardisée exhaustive est mise en place par le gouvernement de Lionel Jospin dans la lignée du vote de sa loi de 1989. Il la présente en détail lors de son discours en discussion générale devant les parlementaires. Il y explicite le constat de départ, issu d'un rapport sur « la réussite à l'école » commandé au recteur Michel Migeon après les résultats jugés catastrophiques de l'évaluation menée par la nouvelle Direction de l'Évaluation et de la Prospective créée par l'ancien ministre de l'éducation René Monory en 1987¹⁰⁸ : pire encore que la moitié des élèves ne maîtrisant pas la lecture en fin de CM2 pour avoir un accès « *autonome à la connaissance* », les résultats indiquent que 20% d'entre eux ne savent absolument pas lire. Cette nouvelle lecture des chiffres, bien que remise en question par de nombreux chercheurs de l'époque¹⁰⁹, amène le recteur à proposer une évaluation annuelle en deux moments (à l'entrée

¹⁰⁸ Selon les conclusions de la DEP publiées dans un article du numéro 14 (janvier-mars 1988) d'Éducation et Formations, les évaluations nationales des élèves en fin de CM2 réalisées en juin 1987 montrent que « à peine plus de 50% de ceux qui ont quitté le CM2 en 1987 maîtrisaient suffisamment la lecture pour pouvoir l'utiliser comme un *moyen d'accès autonome à la connaissance* ».

¹⁰⁹ Voir notamment la critique acerbe publiée dans la revue Pratique, n°63, septembre 1989, p.123-125, qui exprime le « *profond désaccord* » de ses auteurs avec les méthodes, conceptualisations et conclusions du rapport, et fait part sur certains aspects de leur « consternation ».

au CE2 et en sixième) des compétences des élèves en lecture et en écriture. L'objectif serait alors de leur permettre individuellement, selon leurs résultats, de bénéficier d'un accompagnement leur permettant de combler leurs lacunes. Lionel Jospin s'en saisit et intègre cette proposition dans sa loi de 1989 en faisant un dispositif phare de son texte. Il fait, devant les sénateurs, la démonstration de sa nécessité compte tenu des enjeux qu'il juge prioritaire pour l'éducation :

« La priorité absolue, c'est celle des savoirs fondamentaux. Pour donner à tous les enfants la qualité de formation à laquelle ils ont droit, il faut partir de ces savoirs fondamentaux : lire, écrire, calculer, qui sont à la base de toute réussite. Les enfants sont encore trop nombreux, aujourd'hui, à arriver en 6^e avec de graves problèmes de lecture et d'écriture, ainsi que le souligne le recteur Migeon dans le rapport que je lui ai demandé. Nous devons réagir (...).

C'est ce qui fonde le plan d'urgence que je mets en œuvre, dès la rentrée, pour les savoirs fondamentaux. Ce plan s'appuie sur une action sans précédent au sein de l'éducation nationale, c'est-à-dire sur une évaluation systématique des acquis des élèves à l'entrée du cours élémentaire 2^e année et à l'entrée en 6^e. Sans précédent, cette évaluation l'est, tout d'abord, par son ampleur exceptionnelle. Jusqu'à présent, les évaluations du ministère ont porté sur des échantillons. Cette méthode doit, d'ailleurs subsister. Cette fois-ci, 1 720 000 élèves seront concernés.

Cette évaluation est, aussi, exceptionnelle par ses finalités. Elle n'est pas destinée à dresser un constat, ni, encore moins, à venir nourrir les statistiques du ministère. Elle est faite pour agir (...). »

Il remet ainsi, comme il le martèle tout au long de son exposé des motifs, l'enfant au cœur du système éducatif. La « performance » du système éducatif est donc de porter assistance, individuellement et le plus rapidement possible, aux besoins des élèves repérés grâce aux évaluations de début d'année. L'organisation de ces évaluations d'une ampleur inédite et en un temps record est, pour le ministre, « *la démonstration que l'éducation nationale sait, contrairement à ce qui est dit parfois, réagir vite et à grande échelle.* »¹¹⁰

L'évaluation remplit ici un double objectif : elle permet non seulement au niveau des élèves de répondre à des besoins jusqu'ici non repérés, mais elle sert surtout de pilote pour l'action afin de s'assurer que les mêmes problématiques ne

¹¹⁰ Sénat, compte rendus des débats en séance, séance du 27 juin 1989, p.1988, JORF.

se reproduisent plus. Pour le ministre, l'Etat se doit « *d'agir d'une double manière, en mettant en place un véritable plan d'urgence pour tous ceux qui, aujourd'hui, ont des difficultés, mais aussi en organisant plus efficacement l'apprentissage de la lecture à l'école primaire pour offrir à tous les enfants qui viendront ensuite les meilleures chances de réussite* ».

vii. L'évaluation comme outil du projet réformateur

La Loi d'orientation sur l'éducation de Lionel Jospin, votée en 1989, place la notion de projet réformateur au cœur de la définition des actions des établissements d'enseignement, de l'école au lycée, et des actions de l'Etat. L'évaluation est considérée par le ministre comme une véritable « obligation » de l'Etat, dont le rôle est de « déterminer et de garantir les objectifs nationaux ». S'il ne nie pas que des formes d'évaluations pré-existaient déjà, elles ne sont pas suffisantes et ne vont pas dans la direction souhaitée pour répondre aux enjeux de rénovation du système éducatif. L'évaluation doit désormais être globale, concerner tous les échelons et tous les acteurs, dans une démarche constructive et plus seulement de simple bilan. Le ministre l'exprime ainsi devant les sénateurs au moment de l'examen de sa proposition de loi :

« La rénovation du système éducatif en France implique une meilleure évaluation à tous les échelons, réalisée dans un esprit de conseil et de proposition. Il s'agit là d'une autre innovation de cette loi.

L'évaluation s'applique à tout le système éducatif, à savoir les élèves, les étudiants, les personnels, les établissements, les services extérieurs et l'administration centrale. A l'Assemblée nationale, un député avait même ajouté « le ministre ». »

Intéressante précision que celle de ce député reprise par le ministre, qui admet publiquement face à la représentation nationale que l'action des décideurs n'est pas exempte d'évaluation, une vraie nouveauté qui questionne sur les moyens mis en œuvre ensuite pour la concrétiser. L'évaluation de l'action du ministre est-elle celle de la globalité du système éducatif ? Difficile de conclure tant rien ne la définit clairement, ni en 1989 ni maintenant.

Le titre V de la loi, dont le titre est « l'évaluation du système éducatif », composé de trois articles, détaille le rôle et les responsabilités de chacun.

TITRE V. L'évaluation du système éducatif

« Art. 25. - L'inspection générale de l'éducation nationale et l'inspection générale de l'administration de l'éducation nationale procèdent, en liaison avec les services administratifs compétents, à des évaluations départementales, académiques, régionales et nationales qui sont transmises aux présidents et aux rapporteurs des commissions chargées des affaires culturelles du Parlement.

Les évaluations prennent en compte les expériences pédagogiques afin de faire connaître les pratiques innovantes. L'inspection générale de l'éducation nationale et l'inspection générale de l'administration de l'éducation nationale établissent un rapport annuel qui est rendu public.

Le ministre de l'Éducation nationale présente annuellement au Conseil supérieur de l'éducation un rapport sur l'application de la loi. Celui-ci est rendu public.

Art. 26. - Le rapport annuel des établissements publics locaux d'enseignement qui rend compte, notamment, de la mise en œuvre et des résultats du projet d'établissement est transmis au représentant de l'État dans le département, à l'autorité académique et à la collectivité territoriale de rattachement.

Art. 27. - Le comité national d'évaluation des établissements publics à caractère scientifique, culturel et professionnel constitue une autorité administrative indépendante. »

Clairement et en toutes lettres, les inspections générales sont alors désignées comme les responsables de l'évaluation du système éducatif dans sa globalité. L'évaluation du système est bien celle de ses unités d'enseignement à différentes échelles, et la logique d'évaluation standardisée aux indicateurs globaux n'est pas encore celle adoptée par les instances évaluatrices du ministère. Le lien est fait avec les commissions parlementaires chargées des textes relatifs à l'éducation au sein de la commission des affaires culturelles, dans le cadre de leurs missions de contrôle de l'activité gouvernementale inscrite dans la Constitution de la V^e République¹¹¹.

¹¹¹ Philippe Houillon, « Le contrôle extraordinaire du parlement », *Pouvoirs*, vol. 3, n°134, 2010.

La notion de prospective transparait déjà dans la façon d'envisager l'évaluation : elle n'est pas seulement l'outil de contrôle de la bonne utilisation de l'action publique et de ses deniers, mais elle est aussi la lampe de poche qui met en lumière les innovations du quotidien sur le terrain, susceptibles d'essaimer plus largement dans le système éducatif.

Pour cela, au-delà des évaluations statistiques, Lionel Jospin introduit officiellement le monde universitaire dans le panorama des producteurs d'outils d'aide à la décision. Ils ont pour lui un rôle fondamental pour déminer les dossiers complexes, notamment ceux à l'époque au cœur du débat éducatif de l'orientation et du contenu des enseignements et des programmes. En complément de la création du Conseil national des programmes, considérée comme une « instance nécessaire » mais pas suffisante, le ministre souhaite engager une réflexion à la fois académique et issue du terrain en commandant des missions thématiques pilotées par deux pontes du monde de la recherche d'alors, Pierre Bourdieu et François Gros, tous deux professeurs au Collège de France.

Cette réflexion sur les programmes et les contenus d'enseignement n'est pas anodine. Le contenu est construit par ceux qui détiennent le pouvoir de décider de la légitimité du savoir scolaire¹¹², et donc par construction ce qui est donné comme objectif(s) pour l'éducation dans le cadre scolaire (apprentissage de contenus donnés, apprentissage de compétences spécifiques à tous les niveaux, apprentissage de savoir-être et d'attitudes sociales par l'expérience de la vie entre pairs, etc.). Ces objectifs, ces attendus, diffèrent au fur et à mesure du temps en fonction du passage aux plus hautes fonctions décisionnelles d'acteurs aux visions des enjeux parfois diverses. Comme l'énonce Pierre Clément, « *établir un curriculum revient en effet souvent à poser des questions du type : quelle(s) culture(s) scolaire(s) pour quels élèves ? Doit-on enseigner la même chose à tout le monde ? Tous les élèves sont-ils aptes à recevoir et à s'appropriier les mêmes savoirs ? Pour qui faut-il organiser*

¹¹² Pierre Clément, *Réformer les programmes pour changer l'école. Une sociologie historique du champ du pouvoir scolaire*, thèse pour le doctorat de sociologie sous la direction de Bertrand Geay et soutenue à l'Université de Picardie Jules Verne le 11 mars 2013.

le système d'enseignement ? (...) La définition des savoirs scolaires ne peut donc se faire en dehors de toute référence à de tels enjeux, tant elle est inextricablement liée à la hiérarchie scolaire et sociale entre les filières et leurs finalités respectives »¹¹³.

Les questions de contenu des enseignements sont donc un miroir des débats portant sur l'évaluation des résultats de l'éducation. Qu'évalue-t-on et sous quelle forme ? Les acquis des élèves semblent avoir dès les années 1970 la préférence des politiques. On évalue des résultats à un examen donné à un moment donné, construit pour refléter le contenu des programmes enseignés. Avec l'avènement des évaluations standardisées dans les années 1990, la récente Direction de l'Évaluation et de la Prospective devient la main du ministère pour penser et mener l'ensemble des évaluations statistiques sur le territoire national.

2) La statistique comme méthode au centre de la constitution du champ des évaluateurs d'Etat.

Il faut finalement attendre la fin des années 1980 pour que l'Éducation nationale implique officiellement l'ensemble de ses agents dans l'évaluation de son action. L'obligation de moyens n'est plus, l'heure est à l'obligation de résultats. Le ministère redéfinit les actions des agents en termes d'objectifs à atteindre, de moyens à mettre en œuvre et de réalisation des actions dont les fonctionnaires « gestionnaires » ont la charge. Dans une note de service envoyée aux fonctionnaires du ministère de l'Éducation nationale le 12 février 1987 (Note de service n°87-056), on peut ainsi lire que « *chaque gestionnaire à quelque niveau que ce soit (...) [doit] définir les objectifs à atteindre ; prévoir les conditions et les moyens de leur mise en œuvre ; organiser le suivi de la réalisation de ces objectifs ; évaluer les résultats obtenus ; proposer les mesures éventuelles de correction des actions dont il a la charge* ».

Dans la même dynamique, le ministère institutionnalise l'évaluation en son sein en créant la première direction dédiée, chargée officiellement d'apporter

¹¹³ Pierre Clément, *thèse cit.*, p. 22.

au gouvernement les outils d'aide à la décision dont il a besoin pour asseoir ou prévoir son action.

a) L'émergence d'une direction spécifique interne au ministère, la DEP

La fin de la décennie 1980 est cruciale pour saisir la place qu'occupe l'évaluation statistique de l'action éducative ces trente dernières années. Après avoir été sollicitée périodiquement et souvent discrètement par les gouvernements successifs pour préparer ou regarder les effets d'une politique éducative, l'évaluation en éducation est finalement institutionnalisée avec la création d'une direction dédiée en interne au ministère de l'Éducation nationale. Le ministre de l'éducation René Monory crée par décret, en février 1987 la DEP, Direction de l'Evaluation et de la Prospective. Elle succède avantageusement au service de statistiques créé par Claude Seibel. Si elle en reprend certaines méthodes et « l'esprit évaluateur » de ses agents formés à l'école « Seibel », dans la lignée des enseignements de l'ENSAE et des principes d'évaluation de l'INSEE, elle acquiert un nouveau statut qui la rend incontournable pour l'ensemble des acteurs politiques et administratifs du ministère de l'Education nationale.

Les termes choisis pour nommer cette direction sont porteurs de sens. La DEP est une direction à part entière du ministère. Elle n'est plus simple service statistique comme pouvait l'être son ancêtre le service de la prévision, des statistiques et de l'évaluation, mais acquiert le rôle officiel de Direction. Son directeur est nommé en Conseil des ministres par le président de la République. Son organisation ne peut désormais être modifiée que par décret du Premier ministre avec l'accord et la signature du ministre de l'Education nationale. Désormais, les autres directions du ministère devront compter avec une direction ayant une fonction « étude ». Le ministre dote la nouvelle direction de plusieurs missions clés :

« Etablir les éléments d'une prospective de l'éducation débouchant sur des programmes d'action. Pour cela elle assure la production et l'analyse des

informations nécessaires à la conception et à la mise en œuvre de la politique éducative, à la prévision des évolutions du système et de son environnement, enfin à l'évaluation du système éducatif et de la formation. »¹¹⁴

René Monory, homme d'Etat, fils de garagiste ayant gravi pas à pas l'échelle politique et selon sa propre formule « *vulgarisateur d'idées nouvelles* »¹¹⁵, s'est attaché à promouvoir au long de sa carrière l'introduction dans le quotidien des nouvelles technologies, et notamment les NTIC. Il devient ministre de l'Education nationale dans le gouvernement de cohabitation de Jacques Chirac de 1986 à 1988. Il y travaille notamment au renforcement de l'accès à la formation des jeunes. Son passé d'entrepreneur infuse dans sa conception de l'éducation nationale. Il le dit dès l'introduction de son discours de rentrée de septembre 1986¹¹⁶, il souhaite faire de « *l'éducation nationale une grande entreprise moderne dispensant un enseignement de qualité – vérité, efficacité et réalisme* ». Un véritable concentré de terminologies issues du monde de l'entreprise duquel est issu le ministre, attaché par ses affinités personnelles à l'entrée dans la « modernité ».

La vérité, le ministère la doit à la communauté nationale, qui par son travail alloue à l'école le budget le plus important au niveau de l'Etat. Cette priorisation donnée à l'école trouve sa justification dans l'avenir :

« (...) L'école est l'entreprise du futur, elle doit donner aux hommes et aux femmes du troisième millénaire les connaissances et la culture qui leur permettront de s'adapter aux conditions nouvelles de travail et de communication, et de créer un progrès nouveau. »¹¹⁷

L'efficacité et le réalisme souhaités par le ministre, indispensables selon lui pour accompagner les enseignants comme les élèves vers un enseignement de qualité, seront garantis par l'usage systématique d'un outil : l'évaluation. Chargée de mettre la focale sur les difficultés à surmonter, les facteurs de réussite et les enjeux

¹¹⁴ Décret n°87-83 du 10 février 1987 portant organisation de l'administration centrale du ministère de l'Education nationale.

¹¹⁵ Il a notamment créé le parc du Futuroscope de Poitiers.

¹¹⁶ René Monory, ministre de l'Education nationale, « Bonne rentrée », *cahiers de l'éducation nationale*, septembre 1986.

¹¹⁷ *Ibid.*

à venir, l'évaluation est l'outil privilégié pour construire l'avenir du système éducatif. Pour mener à bien son grand plan pour l'éducation, Monory lance une grande réflexion sur l'avenir de l'école qu'il confie à la DEP nouvellement créée. C'est également pour lui le moyen d'assurer la mission léguée par Jean-Pierre Chevènement, ministre de l'Education nationale de 1984 à 1986, qui avait lancé la fameuse formule très appréciée : « amener 80% d'une génération au niveau du baccalauréat »¹¹⁸. Enjeux à venir, financements à prévoir, infrastructures nécessaires, la massification de l'enseignement allait considérablement changer la manière de faire du ministère.

Dans la lignée idéologique du ministre Monory, Jean Pierre Boisivon est chargé par lui de prendre la tête de la nouvelle Direction de l'Evaluation et de la Prospective en 1987, et donc de coordonner ces travaux. Ancien instituteur puis professeur des universités à partir de 1980, secrétaire général de l'Union des Banques de Paris de 1985 à 1897, il correspond bien aux aspirations et aux méthodes de René Monory, au croisement du secteur public et du secteur privé. Pour Boisivon, l'évaluation est nécessaire : c'est bien la structure du système éducatif, à commencer par ses établissements, qui doit s'adapter à ces nouvelles procédures de pilotage qui impliquent, pour atteindre les résultats demandés, d'avoir le plus d'autonomie possible, y compris financière. Il fera plus tard le parallèle entre un établissement et une entreprise, dont le chef d'établissement serait le patron gérant des équipes :

« La notion de projet d'établissement et le développement des procédures d'évaluation imposent une modification de l'organisation des établissements (...). Il faut donner aux chefs d'établissement une réelle capacité d'initiative, dont les conditions ne sont pas encore réunies aujourd'hui... Dans ce système où sont injectés plus de deux cents milliards par an, il n'y a pas un centime qui ne soit affecté ! Les établissements scolaires ont besoin d'une équipe et d'un patron : nulle part je n'ai jamais vu une institution autonome fonctionner s'il n'y a pas un patron entouré d'une équipe. »¹¹⁹

¹¹⁸ Jean-Louis Derouet, « Entre la récupération des savoirs critiques et la construction des standards du management libéral. Recherche, administration et politique en France de 1975 à 2005 », *Revue des sciences de l'éducation*, n°32 (1), 2006.

¹¹⁹ Jean Pierre Boisivon, « L'horizon de l'école », dans *Projet*, n°223, automne 1990.

Jean Pierre Boisivon, s'il enseigne l'économie, n'en est pas pour autant statisticien. Les premières missions confiées par le ministre, pour lesquelles il est recruté à ce poste en premier lieu, n'ont d'ailleurs que peu de lien avec la statistique. Elles ont davantage consisté à réfléchir de façon concertée, avec un ensemble d'acteurs issus à la fois des deux corps d'inspection de l'Education nationale, l'IGEN et l'IGAEN, et des autres directions du ministère, sur les enjeux et les modalités de l'élaboration d'un grand plan pour l'avenir de l'Education nationale, dont le rapport final est présenté par René Monory en décembre 1987. L'introduction des méthodes statistiques dans les travaux de la DEP sera progressive durant les premières années d'existence de la direction et connaître une vraie accélération au tournant des années 1990. Le ministre souhaite lancer une politique ambitieuse d'évaluation, et la direction est attendue pour en être l'opératrice¹²⁰.

La nomination de Claude Thélot à la tête de la DEPP en remplacement de Jean-Pierre Boisivon en 1990 témoigne de cette nouvelle ambition ministérielle. Précédemment responsable de l'évaluation et des statistiques au sein de la direction, anciennement en poste à l'INSEE où il fût responsable de la division Emploi, il impulse, dès sa nomination, une véritable dynamique statistique qui vient conforter l'orientation des missions de la DEP souhaitée par René Monory. L'INSEE devient un lieu de recrutement privilégié pour la DEP qui intègre dans ses équipes des agents issus de l'institut de statistique national. Le rythme des publications s'accélère, qu'elles soient issues d'évaluations menées en interne ou de commandes passées auprès d'acteurs du monde de la recherche. La direction lance plusieurs supports de communication pour populariser ses travaux, les notes d'information, les dossiers de la DEPP, et le recueil d'indicateurs lancé par Claude Thélot dès 1991, l'Etat de l'école. Les équipes sont étoffées, se stabilisant autour de 220 personnes en 1992¹²¹.

¹²⁰ Xavier Pons, « Les statisticiens du ministère de l'Education nationale : évolution d'un métier d'Etat 1957-2007) », *Histoire de l'éducation*, 140-141, 2014.

¹²¹ Xavier Pons, *art. cit.*

c) Un théâtre de confrontation entre plusieurs visions de l'évaluation en éducation : Claude Thélot et Claude Allègre

Le poids de la DEPP dans le pilotage des politiques éducatives a évolué au gré non seulement des volontés ministérielles mais aussi des parcours professionnels antérieurs de ses directeurs et directrices. Ces derniers ont en effet infusé leur vision de la démarche évaluative dans le pilotage de l'action publique en éducation. Par sa longévité à la tête de la DEPP et par la mue de la Direction qu'il a impulsée, Claude Thélot reste largement cité comme étant le façonneur de la DEPP telle qu'elle existe encore près de trente ans plus tard. Il est celui qui a su combiner à la fois sa propre vision pour la DEPP avec celle, prospective, initiée par Jean-Pierre Boisivon pendant trois années et celle plus ancienne, statistique, de Claude Seibel. Christian Forestier, figure à la fois administrative et politique de l'Education nationale dès le premier mandat de François Mitterrand de 1981, initiateur du Haut Conseil de l'Education en 2003, dira ainsi de lui qu'il est celui à qui « *revient le mérite de la mise en place d'une grande politique d'évaluation des performances des élèves et des étudiants.* »¹²²

Thélot reviendra plus tard sur ce qui constituait pour lui alors l'évaluation. Au milieu des années 1990, l'évaluation du système éducatif avait principalement deux objectifs, celui d'évaluer l'état du système et des unités d'enseignement, et de fournir des outils d'évaluation aux acteurs pour qu'ils s'en servent comme levier d'action. L'évaluation n'est pas encore celle de la « performance » du système par rapport à des objectifs préétablis. Elle n'est pas non plus celle des manques ou des faiblesses du système en prévision d'une réforme déjà réfléchie par les décideurs politiques. Les indicateurs sont pensés au moment de la mise en œuvre du processus évaluatif, en fonction de l'environnement dans lequel il s'inscrit. La structure de l'évaluation est alors un triptyque : coût, fonctionnement (ce que Thélot appelle les « services offerts »), et résultats. Libre à l'évaluateur ensuite d'associer ces dimensions pour obtenir au choix efficacité, efficience (en

¹²² Christian Forestier, « Le meilleur système du monde ? » dans Jean-Claude Emin, Christian Forestier, Claude Thélot, *Que vaut l'enseignement en France ?*, Editions Stock, 2007, p.1.

d'autres termes son efficacité rapportée à son coût), ou rendement du système (ses résultats rapportés à son coût). C'est cette troisième option qui remporte l'adhésion de Thélot, en ce qu'elle prend en compte ce qui légitime réellement la politique éducative, ce qu'elle permet d'atteindre. L'efficacité n'est pour lui pas une dimension pertinente lorsqu'il s'agit de politique éducative, car elle évacue cette légitimité en prenant indûment les services rendus pour les objectifs finaux de l'action publique en éducation.

Ces évaluations, qui sont pour Thélot une véritable nécessité au regard du devoir de transparence du service public aux citoyens qui le financent, permettent ainsi de donner un état des lieux de la « qualité » du système. Par leur centrage autant que possible sur le rendement de l'éducation, elles nécessitent que les objectifs assignés à l'éducation soient clairement établis. C'est là selon Thélot le vrai point faible des réflexions sur l'éducation en France. Elle est

« à la fois trop et pas assez développée : elle l'est trop, en ce sens que nulle part ailleurs dans le monde on ne s'interroge autant, aussi souvent et en des termes aussi généraux sur ce que doit réussir l'école ; elle ne l'est pas assez, en ce sens que la réflexion, pour être consensuelle, reste trop abstraite, trop générale pour servir réellement de guide aux acteurs ».

Le guide de l'action publique, voilà bien un objectif majeur pour Thélot qui développe à la DEPP une véritable culture de l'évaluation se fondant largement sur son expérience de statisticien. Personnalité aux compétences reconnues par les responsables politiques, et notamment les ministres de l'éducation nationale successifs Lionel Jospin, Jack Lang et François Bayrou, Claude Thélot fera des évaluations menées par celle qui est encore la DEP de véritables outils aux mains des décideurs politiques.

La direction développe sa boîte à outils statistiques et prend en charge dès la rentrée scolaire de 1989 les évaluations voulues par Lionel Jospin en CE2 et en 6^e. Elle les mène sur le terrain bien sûr, mais les conçoit également, principalement comme des évaluations uniquement diagnostic – légèrement en décalage donc de l'objectif de pilotage souhaité par le ministre au moment des débats parlementaires.

La DEP reprend les méthodologies de projection notamment, mises en place dès les années 1970 pour prévoir à 10 ans les besoins en moyens financiers et humains dans les académies et préparer les rentrées scolaires¹²³. Ces projections ont toujours reposé sur la constitution préalable d'un ensemble de paramètres clés stratégiques, liés aux décisions gouvernementales en matière de politique éducative (par exemple l'orientation des élèves en fin de collège, l'évolution du taux de réussite aux examens ou du redoublement,...). Les travaux de l'ancien service statistique et de la nouvelle direction de l'évaluation sont indissociables des objectifs déterminés par les cabinets ministériels. Cabinet et DEP travaillent en grande proximité, au moins pour déterminer les grands axes stratégiques de l'action à évaluer.

En 1992, lorsque Jack Lang devient ministre de l'Education nationale, le travail main dans la main de son cabinet avec la DEP s'inscrit dans la continuité des travaux précédents. Le ministre apprécie Claude Thélot – qui se définira d'ailleurs dans un écrit comme « son » directeur de l'Evaluation et de la Prospective¹²⁴ - et lui confie la mission de rédiger un premier rapport sur le niveau des élèves qui lui servira de tribune à l'Assemblée Nationale. La DEP se voit confier la responsabilité par le ministre de mener l'ensemble des évaluations diagnostique à l'entrée en seconde la même année, et relance les évaluations bilans au milieu des années 1990. L'arrivée dans les années 1990 des techniques informatiques permet de doter la direction d'outils performants adaptés à ses besoins de gagner en efficacité.¹²⁵

Jusqu'à l'arrivée de Claude Allègre à la tête du ministère en 1998, la DEP est un évaluateur légitime par sa maîtrise des statistiques appliquées à l'évaluation des résultats de l'éducation, encore largement vue par le politique sous le prisme de l'évaluation des résultats individuels des élèves.

¹²³ Jean-Richard Cytermann, « Evaluer et prévoir. De la DEP à la DPD », *Perspectives documentaires en éducation*, n°54, 2001.

¹²⁴ Jean-Richard Cytermann, *art. cit.*

¹²⁵ Xavier Pons, « Les statisticiens du ministère de l'Education nationale : évolution d'un métier d'Etat 1957-2007 », *Histoire de l'éducation*, 140-141, 2014.

d) L'indépendance des évaluateurs en question

En 1998, Claude Allègre devient ministre de l'Éducation nationale et le crédit, la portée des évaluations menées par la DEP changent du tout au tout. Pour le ministre, l'administration observatoire doit être différenciée du service d'analyse statistique interne au ministère. Dans sa volonté de se lancer dans le courant d'internationalisation concurrentielle de l'éducation, le ministre se sépare également de ce qui représente un risque de blocage de son action au nom de corporatismes peu enclins à s'ouvrir au « marché » de l'éducation¹²⁶.

Le ministre va jusqu'à changer le nom de la direction en actant la disparition de l'évaluation des travaux internes au ministère. Elle devient la DPD, Direction de la programmation et du développement. Le vocabulaire est significatif. La portée prospective des évaluations menées par la DEP, au cœur de ses missions, est transformée en une fonction programmatrice qui n'était pas la sienne. Thélot démissionne suite à l'arrivée du nouveau ministre. Ni les missions ni les méthodes ne correspondent à celles développées ces dix dernières années.

Nombreuses sont notamment à l'époque les voix des agents de l'ancienne DEP à s'élever contre ce qui est vécu comme des pressions de la part du ministre pour faire correspondre les chiffres des évaluations avec ceux souhaités par Allègre pour conforter ses choix politiques. La journaliste Anne Fohr rapporte des témoignages de statisticiens dans un article du *Nouvel Observateur* :

« Une censure déguisée et des pressions discrètes sont apparues ; des rapports pourtant commandés par les cabinets des ministres sont passés à la trappe, et des auteurs ont dû corriger leurs publications [...]. On aurait également demandé au service de plancher sur les fameux 12% d'absentéisme enseignant avancé par le ministre Claude Allègre : « nous avons additionné des choux et des carottes » plaisante un statisticien, « sans succès évidemment » [...]. Il y eut aussi une étude bâclée sur la violence dans les établissements scolaires, réalisée et présentée aux recteurs par une béotienne zélée : les jours de grève des enseignants y étaient comptabilisés avec le racket, et les résultats assez désopilants pour que le directeur de cabinet du ministre ait ordonné sa disparition [...]. Au total, au-delà de

¹²⁶ Jean-Louis Derouet, « Entre la récupération des savoirs critiques et la construction des standards du management libéral. Recherche, administration et politique en France de 1975 à 2005 », *Revue des sciences de l'éducation*, 32 (1), 2006.

l'inquiétude et de l'amertume, le personnel conteste la conception « instrumentale » que Claude Allègre a des chiffres et de leur évaluation. Leitmotiv : « il confond l'information publique et la communication politique ». »¹²⁷

Claude Thélot n'aura de cesse de défendre l'indépendance de la DEP disparue, condition sine qua non selon lui de la validité de ses études et qu'il avait pris soin d'assurer lorsqu'il en avait la direction. Il définit trois critères nécessaires de qualité, tous remplis par la DEP. Le premier est celui de la maîtrise de son programme de travail, assurée à l'époque par la publication en début d'année du programme prévu pour l'année à venir, sans que le ministre ne puisse interférer dans les sujets d'étude retenus. Le second est celui de la compétence des évaluateurs, assurée selon lui, on l'imagine, par ses anciens collaborateurs de la DEP issus de l'INSEE et des personnels à la fois de terrain et qualifiés de l'Education nationale.

L'indépendance de l'organe évaluateur n'a donc a priori, selon cette définition, rien à voir avec son statut ou son positionnement hiérarchique vis-à-vis du ministre. En ce sens, la DEP devenue DPD aurait très bien pu tenir lieu d'évaluateur indépendant des actions et réformes entreprises par le ministère, au côté de tout l'arsenal évaluateur existant alors, et notamment les universitaires et évaluateurs externes au ministère régulièrement sollicités par le cabinet de Claude Allègre. Pour Claude Thélot, la disparition de la DEP telle qu'il l'avait pensée est en réalité liée au manque d'intérêt du ministre pour l'évaluation. A mots – pas tout à fait – couverts, il dira ainsi à la fin des années 2000 :

« Une direction d'administration centrale comme l'était la DPD pouvait parfaitement remplir ces trois critères, même si elle faisait partie du ministère de l'Education nationale : il faut (et il suffit) que le ministre, conscient de l'importance de l'évaluation et désireux qu'elle se développe, respecte ces trois conditions. Dès lors que l'on voulait conduire une grande politique d'évaluation, le respect de ces conditions me paraissait s'imposer ».

¹²⁷ Anne Fohr, *Le thermomètre de l'école est bloqué*, Nouvel observateur, avril 1999.

Dès le départ de Claude Allègre du ministère en 2000, son successeur Jack Lang reprend contact avec Claude Thélot, symbolisant ainsi son positionnement sur sa « ligne » en matière d'évaluation des politiques éducatives. Il relance les travaux de la DPD. Il lui commande, par la bouche de Christian Forestier alors directeur de cabinet, de travailler à la construction d'une instance chargée de « *dire la vérité aux Français et à leurs dirigeants sur les problèmes de notre école* ¹²⁸», en mutualisant notamment les ressources en termes d'évaluation des résultats du système éducatif et les travaux de recherche nationaux et internationaux sur le sujet. Le Haut Conseil de l'Education naît, et sa présidence est immédiatement proposée à Claude Thélot qui l'accepte. Si le Haut Conseil n'est pas tout à fait en lien direct avec la DEP, sa création et la nomination à sa tête entérinent la volonté ministérielle de redonner aux évaluations nationales et à ses évaluateurs un rôle prépondérant dans les réflexions sur les besoins d'évolution du système éducatif. Les évaluations menées par la DPD retrouvent un lectorat politique. Cette revalorisation de ses missions aboutira à refondre ses missions et sa dénomination pour rendre compte de ce nouveau statut. La DPD devient en 2006 la DEPP, Direction de l'évaluation, de la prospective et de la performance.

e) La statistique comme méthode de l'évaluation de l'action publique

La DEPP voit ses missions élargies à sa création. Elle est chargée, en plus de ses missions précédentes, d'évaluer spécifiquement la performance du système éducatif français, principalement d'un point de vue statistique. Son rôle a bien évolué depuis sa création en 1987. Alors qu'à sa naissance elle avait été chargée par René Monory de coordonner et prendre part à une réflexion prospective d'ampleur sur l'avenir du système éducatif français, sans véritable focus unique sur l'utilisation des outils statistiques, elle est peu à peu en charge de l'ensemble des évaluations standardisées menées par les gouvernements successifs. A l'ère

¹²⁸ Christian Forestier, « Le meilleur système du monde ? » dans Jean-Claude Emin, Christian Forestier, Claude Thélot, *Que vaut l'enseignement en France ?*, Editions Stock, 2001.

numérique, le développement des outils informatiques permet de développer des tests de plus grande ampleur plus facilement, y compris en les faisant administrer directement sur numérique dans les établissements.

En 2009, la DEPP est intégrée à la liste des services et directions statistiques ministériels pour le ministère de l'éducation nationale et de la jeunesse. Elle rejoint la sous-direction des systèmes d'information et des études statistiques (SIES) qui intervient sur le champ de compétence du ministère récemment créé de l'Enseignement Supérieur, de la Recherche et de l'Innovation. Elle s'engage par cette intégration à respecter un certain nombre de principes fondamentaux permettant de garantir le sérieux de ses publications et de ses résultats, ainsi que l'objectivité et la neutralité indispensables à leur utilisation par le politique.

A partir de 2014, la DEPP renforce son autonomie pour renforcer dans le même mouvement son indépendance vis-à-vis du politique : son programme d'activité est rendu public, qui comprend nombre d'études parfois en coopération avec le monde universitaire. Cette relation est notamment inscrite, au milieu de toutes les autres missions allouées à la DEPP, dans l'article 6 du décret n°2014-133 du 17 février 2014 qui fixe l'organisation de l'administration centrale des deux ministères :

« IV – La direction de l'évaluation, de la prospective et de la performance rend compte de l'état du système de formation et d'éducation au moyen d'études qu'elle mène et de recherches qu'elle engage avec des établissements d'enseignement supérieur ou des organismes de recherche. Par la diffusion de l'ensemble de ses travaux et productions, elle contribue à éclairer le débat public sur l'éducation et la formation. »

Son rôle d'expert au service du pilotage de l'ensemble de l'action du ministère de l'Éducation nationale est écrit noir sur blanc dans ce décret. Ses liens avec la Direction Générale de l'Enseignement Scolaire (Dgesc) sont toutefois durablement tenus, en particulier lorsqu'il s'agit de préparer son programme d'étude annuel et de réfléchir aux publications à venir. Cette distanciation de la Dgesc sera un vrai sujet d'interrogation pour les observateurs et évaluateurs de l'évaluation du système éducatif français (voir supra). En creux se lisent les

prérogatives que chacun s'alloue et les conflits que cela peut créer lorsque les missions théoriques et les missions en pratique diffèrent du fait de vents contraires initiés par les autres acteurs de l'évaluation.

En termes de positionnement méthodologique, à la fin de la période, la DEPP est en charge de toutes les évaluations statistiques standardisées ou non menées à tous les niveaux scolaires sur le territoire national. Elle est chargée de concevoir et de gérer toute l'information statistique mise au service du ministère, et est officiellement l'interlocutrice des organisations internationales, en représentation du ministère de l'Éducation nationale, pour tous les projets européens ou internationaux évaluant les performances des systèmes éducatifs. Finalement, son rôle de conseil à proprement parler, toujours bien présent dans les textes qui définissent ses missions, n'a plus rien à avoir avec celui qui était le sien à sa création en 1987. L'évaluation de la performance du système éducatif, s'il est lié à des réflexions prospectives, se place résolument dans le champ de la statistique et de l'analyse à plat des données recueillies, sans que cela n'aboutisse officiellement à la publication de recommandations de pilotage, laissées théoriquement aux mains de la Dgesco. En pratique, la DEPP parvient tout de même à amener certains sujets qu'elle considère (ses directeurs en tout cas) comme prioritaires, en publiant des notes d'information sur des sujets précis pour créer un coup de projecteur présenté aux cabinets. Elle travaille ainsi par exemple spécifiquement sur les questions de mixité sociale, d'éducation prioritaire¹²⁹ ou encore de redoublement, pour induire, renforcer ou éclairer les débats en interne du ministère ou le débat public sur des sujets qui passionnent souvent chercheurs, syndicats de l'éducation nationale, pédagogues.

Toutefois, malgré ses missions d'évaluation centrales, la DEPP loin d'avoir le monopole des évaluations de l'action éducative, est rapidement entourée de toute une série d'institutions et d'acteurs qui entrent progressivement sur le même champ d'action et produisent des contenus évaluatifs d'aide à la décision. Si en

¹²⁹ Mission sur l'Évaluation et l'organisation de la fonction statistique au ministère de l'Éducation nationale, de l'Enseignement Supérieur et de la Recherche (MENESR), *Évaluation et organisation de la fonction statistique au ministère de l'Éducation nationale, de l'Enseignement Supérieur et de la Recherche*, mai 2016, p. 6.

théorie chacun a son terrain d'action prioritaire, en pratique les frontières entre leurs rôles respectifs ne sont parfois pas évidentes, y compris pour les ministres de l'éducation successifs qui peuvent ajouter de nouveaux évaluateurs au gré des grands textes législatifs qu'ils portent.

f) L'évaluation tous azimuts : acteurs et « esprit » de l'évaluation

A la fin des années 2010, le théâtre de l'évaluation de l'action éducative est occupé par une pléthore d'acteurs aux centres d'intérêt et aux zones d'influence différentes. Cette diversité permet sans doute un croisement des regards sur les réponses apportées aux enjeux de l'action éducative, garantissant ainsi aux décideurs politiques qui interviennent parfois en dehors de leur strict champ d'attribution. Elle rend toutefois floue les résultats des évaluations menées en termes d'implication de politique éducative¹³⁰, en ne donnant pas aux décideurs un outil « clé en main » leur permettant d'orienter leurs réformes. Les évaluations nécessitent confrontation, analyse croisée, formulation d'hypothèses d'actions qui ne sont finalement à la charge d'aucun organe évaluateur, ou du moins d'aucun n'ayant à rendre compte de ses conclusions dans une démarche de pilotage de l'action publique à la suite d'une commande politique. Ce manque est d'ailleurs résumé de façon acerbe par la Cour des Comptes qui estime, dans un rapport fourni publié en 2003 sur la gestion du système éducatif – et en centrant clairement son propos comme l'indique son titre dans une perspective de la recherche de pistes de réformes administratives¹³¹ – que :

« Les réponses apportées par les travaux d'évaluation du système scolaire restent encore limitées. Ces travaux ne permettent pas d'établir une relation entre les moyens et les résultats de l'école. Malgré la richesse et la variété des données statistiques qu'ils utilisent, ils ne parviennent pas encore à identifier les facteurs à l'œuvre dans le

¹³⁰ Jean-Richard Cytermann, « La contribution des outils statistiques et d'évaluation à la modernisation de l'Éducation nationale », *Politiques et management public*, vol.23, n°1, 2005.

¹³¹ La Cour des Comptes définit ainsi l'objectif de son rapport : « Identifier ses objectifs (du système actuel), appréhender ses moyens et déterminer si ses résultats sont précisément évalués pour améliorer sa gestion a ainsi été le guide de la démarche ». Cour des Comptes, *La gestion du système éducatif*, avril 2003, p. 8.

processus d'amélioration de la formation des élèves. Enfin leur utilisation à des fins de décision reste insuffisante. »¹³²

Deux types d'évaluateurs produisent en particulier des évaluations du système éducatif¹³³ pouvant venir s'entrechoquer avec les missions officielles de la DEPP, les instances administratives et créées à l'initiative des ministres de l'Education nationale, et les acteurs du monde de la recherche.

i. Dgesco et conseils d'évaluation de l'éducation : flou des missions, flou du positionnement

Si la Dgesco, Direction générale de l'enseignement scolaire, a théoriquement pour missions de gérer les questions de performance budgétaires du ministère, cela implique qu'elle a intégré dans ses missions l'évaluation prospective des besoins du système éducatif. La Direction a pour mission principale, selon ses propres mots, d'élaborer la politique éducative et de mettre en œuvre les programmes relatifs aux écoles, collèges, lycées d'enseignement général et professionnel. En ce sens elle doit travailler main dans la main avec l'Inspection Générale de l'Education Nationale, puisqu'elle est officiellement l'instance prescriptrice des évaluations des élèves. Elle est également en lien étroit avec la DEPP dans ce même cadre. Au-delà du suivi des résultats des évaluations des résultats des élèves, elle peut également bénéficier des expérimentations et évaluations menées par la DEPP pour construire ou faire évoluer les politiques éducatives. D'autant plus que si l'on regarde dans le détail la description de ses champs d'intervention, la Dgesco intervient peu dans la démarche évaluative du système éducatif mais doit bien définir le cadre d'action du gouvernement sur

¹³²*Ibid.*, p.110.

¹³³ Nous ne nous attarderons pas sur les missions et le rôle de l'IGEN et de l'IGAEN qui ne se positionnent pas du côté de l'évaluation du système éducatif dans sa globalité et à des fins de pilotage. Ce sont des inspections chargées d'évaluer les personnels de l'Éducation nationale, dont l'activité n'est pas orientée vers l'apport de données relatives au système éducatif dans son ensemble. Ils ne se sont prononcés que rarement sur les grandes orientations politiques choisies par les différents gouvernements en matière éducative à travers des rapports d'évaluation portant sur les concepts stratégiques de l'action éducative.

différents sujets centraux : la formation des personnels enseignants, la réglementation de l'organisation et du fonctionnement des écoles et établissements scolaires du second degré, la politique de scolarisation des élèves en situation de handicap, la politique de l'éducation prioritaire, la politique de vie scolaire, la réponse aux questions pédagogiques posées dans le cadre de l'enseignement privé, entre autres.

Mais finalement la DEPP et la Dgesco travaillent en parallèle sans véritable concertation ou action conjointe. Ces trajectoires non sécantes, étonnantes compte tenu de leurs compétences respectives, sont soulignées au cours d'évaluations externes au ministère, réalisées notamment par la Cour des Comptes et la Mission d'évaluation. Toutes souligneront ce qui sera qualifié avec diplomatie de « divergences d'appréciations » entre les deux directions sur « *l'implication attendue concernant certains outils d'aide au pilotage ou l'évaluation de nouveaux dispositifs* »¹³⁴. Sans que cela ne soit jamais réellement tranché au moment des diverses réorganisations du ministère de l'Éducation nationale¹³⁵.

La Dgesco se place elle volontiers à l'interface des évaluateurs (DEPP, mais aussi inspections générales) et des cabinets ministériels. Elle est dotée d'un département en charge notamment de suivre les expérimentations menées par la DEPP sur demande du ministre, le département de la recherche et du développement, de l'innovation et de l'expérimentation.

Les missions de la Dgesco et du département en question sont ainsi résumées par l'une des responsables :

« Le cabinet arrive avec un projet bien sûr, et avec une vision éclairée de ce qu'il convient de faire sur le champ de l'égalité, pour améliorer le collège ou que sais-je. Et puis ensuite effectivement les directions instruisent la procédure qui leur est proposée de suivre. C'est la DEPP qui est responsable des évaluations. Alors la DEPP mène des évaluations qui sont effectivement quantitatives et qualitatives. Elle évalue aussi les résultats des expérimentations sans aucun doute, mais les évaluations sont aussi confiées

¹³⁴ MENESR, *op.cit.*, p.6.

¹³⁵ Notamment celle de 2014, décrite dans le décret n°2014-133 du 17 février qui fixe l'organisation de l'administration centrale des ministères de l'Éducation nationale et de l'enseignement supérieur et de la recherche, qui précise les missions de la DEPP notamment dans son article 6.

aux inspections générales. Donc on a une évaluation qui est répartie sur un grand nombre de têtes, un grand nombre d'instances plutôt, dans notre ministère. Et finalement, nous, notre travail c'est de prendre connaissance de ces évaluations pour en faire notre miel et travailler avec ça. »¹³⁶

Du côté de la DEPP, ce rôle d'analyste des résultats d'évaluation de la Dgesco semble aller de soi, même si les publications des résultats donnent souvent lieu à des recommandations de fait sous la forme de « focus » sur des sujets jugés prioritaires et stratégiques :

« La DEPP ne fait pas de recommandations à partir des résultats, elle met en lumière les chiffres. L'interprétation se fait au niveau de la Dgesco ou du cabinet du ministère. Tout de même, la DEPP avait par exemple depuis longtemps insisté sur les inégalités sociales reproduites voire aggravées par le système éducatif. »¹³⁷

La frontière entre « mise en lumière » appuyée au fil des rapports et recommandations est bien floue. Son rôle d'abeille de l'évaluation en éducation se mélange parfois par ailleurs avec celui d'instances créées au gré de textes législatifs réformateurs. Différents hauts conseils ou comités ad hoc sont en effet créés par plusieurs textes réformant le système éducatif et viennent s'ajouter à la liste des observateurs évaluateurs du pilotage de l'action éducative.

On y trouve notamment le Haut Conseil de l'évaluation de l'école souhaité par Jack Lang qui est constitué en 2000 par décret. Le gouvernement lui avait assigné un rôle pluriel, finalement presque au croisement de celui de la DEPP et de la Dgesco. Le Haut Conseil était en effet chargé d'évaluer les évaluations menées par le ministère tout en en tirant les conclusions les plus structurantes, et de proposer des évaluations complémentaires sur des sujets particuliers. Composé de membres tous nommés par le ministre de l'Éducation nationale, il est remplacé à l'occasion de la loi d'orientation et de programme pour l'avenir de l'école du 23 avril 2005 par le Haut Conseil de l'Éducation, à la composition plus variée et indépendante en théorie du ministre. Il est cette fois chargé de la mission

¹³⁶ Entretien avec Isabelle Robin, Directrice du département de la recherche et du développement, de l'innovation et de l'expérimentation, Dgesco, 7 avril 2017.

¹³⁷ Entretien avec Bruno Trosseille, Directeur du bureau des évaluations des élèves, DEPP, 5 janvier 2016.

pas plus précise que son prédécesseur d'« adresser chaque année un bilan des résultats obtenus par le système éducatif », y compris des expérimentations menées dans les écoles et établissements du second degré. Soit, à peu de choses près l'une des missions de la Dgesco. Le HCE tente toutefois d'appuyer la création d'une instance évaluatrice d'évaluations indépendante, pouvant s'autosaisir et ayant le statut d'autorité administrative indépendante. Ce statut, qui aurait pu donner aux travaux du HCE une teinte différente de ceux de la Dgesco bien qu'orientés vers une direction proche, ne sera jamais accordé sans que les raisons de cet immobilisme ne soient réellement explicitées. Il est remplacé en 2013 par le Conseil supérieur des programmes à l'occasion de la réforme menée par Vincent Peillon pour la refondation de l'École de la République, perdant alors sa mission d'observateur et d'analyste des résultats du système éducatif, pour se concentrer sur les enjeux de transparence de la construction des programmes d'enseignements et des épreuves d'examen des élèves et de concours enseignants.

ii. La sollicitation discrète du monde universitaire

Face à des objets d'évaluations variés (résultats des élèves, résultats des établissements, effets des politiques menées, prévisions des besoins en moyens humains et financiers, etc.) de nombreux acteurs sont mobilisés pour mener des évaluations avec des façons singulières de penser les problématiques et les méthodes d'évaluation. Les nouveaux acteurs institutionnels ne sont pas les seuls à être sollicités par les différents ministres de l'éducation dans la lignée de l'envie de Lionel Jospin d'intégrer des regards pluriels sur l'état du système éducatif et ses orientations.

Parmi les évaluateurs à l'œuvre à partir des Trente Glorieuses fournissant des outils de pilotage public, l'université française fait figure discrète. Elle ne s'est pas particulièrement intéressée aux politiques éducatives dans une visée opérationnelle et évaluative avant les années 1970, années où apparaît un nouveau courant de recherche dans la lignée des travaux anglo-saxons de la *school*

effectiveness. De plus en plus de publications portent sur les politiques éducatives nationales et sur l'efficacité de l'enseignement¹³⁸. La science de l'éducation et la sociologie semblent encore restées au carrefour d'une ambivalence de fond : si elles sont réticentes à adopter une démarche évaluative, l'étude du système éducatif et des effets des actions politiques entreprises sont des objets au cœur de leurs prérogatives et de leurs intérêts.

Si les travaux nombreux publiés sur ces objets d'étude n'ont pas pour objectif premier de servir d'outil de pilotage aux décideurs politiques, le monde de la recherche est sollicité, particulièrement à partir des années 1980¹³⁹, par les gouvernements successifs pour travailler à des évaluations commandées, y compris après la création de la DEP. Les travaux sont confiés à des experts externes au ministère issus pour beaucoup du monde universitaire. Souvent passés presque inaperçus au commencement, ils acquièrent progressivement une popularité qui les posent en interlocuteurs légitimes des décideurs politiques pour orienter leur action. En 1983, Antoine Prost travaille aux orientations sur les seconds cycles des lycées, et en remet le rapport à Alain Savary, alors ministre de l'Éducation nationale. Bien que loué par la communauté éducative pour les préconisations qu'il fait et la méthode avec laquelle il a été constitué, il ne sera pas réellement utilisé à des fins de pilotage. Quelques années plus tard, un important travail est mené par la Commission Bourdieu Gros en 1998, en amont de la réforme menée par Lionel Jospin, à la suite d'un rapport critique remis par Pierre Bourdieu au ministre portant des critiques assez vives sur le contenu des programmes scolaires notamment et leur manque de lien avec les enjeux sociaux et scientifiques à venir. On peut également citer les rapports ayant servi de fondement aux réformes du collège menées par Ségolène Royal en 1999, alors ministre déléguée chargée de l'enseignement scolaire, et Jack Lang en 2001, ministre de l'Éducation nationale. La première s'appuie largement sur le rapport

¹³⁸ Pierre Merle, « L'efficacité de l'enseignement », *Revue française de sociologie*, vol. 39, n° 3, 1998.

¹³⁹ Xavier Pons, *Évaluer l'action éducative, des professionnels en concurrence*, Education et société, PUF, Paris, 2010.

remis par les sociologues François Dubet¹⁴⁰ et Marie Duru-Bellat¹⁴¹ et l'historien et inspecteur général de l'Éducation nationale Alain Bergounioux¹⁴² intitulé *Le Collège de l'An 2000*. Le second travaille sur l'adaptation des pratiques pédagogiques sans toucher à l'architecture du collège à partir des recommandations de l'historien Philippe Joutard¹⁴³ formulées dans son « *Rapport sur l'évolution du collège* ». Les deux experts, bien que faisant carrière dans le monde universitaire, n'en sont pas moins proches des décideurs politiques et se sont investis à divers moments pour apporter leur regard évaluatif *ex post* ou *ex ante* sur les politiques éducatives.

La sollicitation ministérielle du monde de la recherche pour produire des évaluations du système éducatif ad hoc ne prend pas particulièrement d'ampleur dans les années qui suivent l'an 2000. Ils sont en revanche régulièrement incités à participer aux instances d'évaluation créées pour donner une analyse et des préconisations d'ensemble des publications et évaluations, tous acteurs confondus, portant sur les résultats du système éducatif français. Le Haut Conseil

¹⁴⁰ François Dubet a enseigné la sociologie à l'Université Bordeaux II jusqu'en 2013. Il fut membre jusqu'en 2010 d'un groupe d'experts mandaté par le ministre de l'Éducation nationale Luc Chatel pour réfléchir à la refonte du programme de sciences économiques et sociales en amont de la réforme du lycée qu'il voulait mener. Il en démissionne avec fracas à la publication du programme. Très critique d'un système scolaire qu'il juge inégalitaire et excluant, François Dubet a appelé publiquement à voter pour Ségolène Royal, candidate de la gauche socialiste à l'élection présidentielle de 2007.

¹⁴¹ Sociologue spécialiste de l'éducation et de l'enseignement supérieur, Marie Duru-Bellat est professeur de sociologie à Science Po et chercheuse à l'Observatoire sociologique du changement du CNRS. Elle s'est investie dans des instances consultatives en éducation, et notamment au Haut Conseil pour l'évaluation de l'école à partir de 2000 et la Commission du débat national Université Emploi en 2006.

¹⁴² Alain Bergounioux est un « universitaire expert » actif au sein de plusieurs cabinets ministériels dès la fin des années 1980. Il conseille Michel Rocard, Premier ministre, puis travaille aux côtés de Catherine Trautmann, ministre de la Culture et de la Communication. Il conseille enfin Michel Sapin au ministère de la Fonction publique et de la Réforme de l'État au début des années 2000. Une décennie après, en 2012, Vincent Peillon lui confie, avec Rémy Schwartz et Laurence Loeffel, une mission portant sur la morale laïque à l'école en prévision de la mise en place de l'enseignement moral et civique à la rentrée scolaire de 2015.

¹⁴³ Philippe Joutard est professeur émérite d'histoire moderne à l'université de Provence Aix-Marseille I et recteur d'académie honoraire. Il fut recteur d'académie de Besançon de 1989 à 1992 et de Toulouse de 1992 à 1997. Il avait présidé de 1988 à 1989 une commission ministérielle sur l'enseignement de l'histoire géographie et a travaillé, pour le ministère de l'Éducation nationale et de la Culture, sur un rapport portant sur la Grande pauvreté et la réussite scolaire (publié en octobre 1992). Il est donc un habitué des travaux d'expertise pour les différents ministres de l'Éducation nationale.

de l'éducation, comportera parmi ses membres des professeurs d'université et des directeurs de recherche à l'Inserm. Le Conseil supérieur des programmes qui lui succède comportera parmi ses membres des personnalités qualifiées, dont plusieurs enseignants chercheurs. Le Centre national d'évaluation du système scolaire a quant à lui à sa tête Nathalie Mons, sociologue spécialiste des politiques éducatives françaises et étrangères.

Cette pluralité de modalités de contribution du monde de la recherche aux souhaits des ministres successifs d'avoir une vision d'ensemble des résultats des évaluations sur le fonctionnement du système éducatif a deux effets sur lesquels nous reviendrons plus en détail au chapitre 4 de cette thèse. D'une part, ils font du monde de la recherche un interlocuteur clé pour les agents de la DEPP au moment où ils souhaitent légitimer la valeur des évaluations qu'ils mènent pour montrer un certain nombre de problèmes publics sur lesquels agir pour améliorer le fonctionnement du système. D'autre part, ils contribuent à créer un foisonnement de producteurs d'évaluations de l'action éducative pour l'action gouvernementale, qui tend à conférer aux travaux de la DEPP davantage un rôle de pourvoyeur de données descriptives analysées par d'autres, que de réelle direction de l'évaluation à visée prospective mettant en lumière des orientations souhaitables de l'action éducative.

Conclusion du chapitre

Au début du XXI^e siècle, l'évaluation des politiques éducatives s'est constituée en quelques décennies en un véritable champ d'experts structurés, internes au ministère de l'Éducation nationale pour la plupart, dont le nombre et les missions se sont considérablement élargis à partir des années 1980. L'outil évaluateur s'est progressivement dédoublé : n'étant plus seulement le moyen pour les citoyens d'évaluer l'ampleur des missions d'intérêt général remplies par l'État, elle est aussi la lampe qui éclaire les décisions des décideurs publics, afin de les accompagner dans leurs choix et les aider à améliorer l'efficacité des

politiques publiques qu'ils conduisent¹⁴⁴. Cette évolution s'inscrit dans le cadre des politiques de réforme de l'administration revendiquées depuis les années 1990 sous le nom de « réforme de l'État », et qui se veulent corriger les dysfonctionnements « objectifs » de l'État¹⁴⁵. L'État et ses administrations rendent alors visibles les défaillances de l'appareil d'État et érigent la réforme en objectif de leur action. La sociologie, l'économie, la gestion, le management, sont alors autant de disciplines qui rendent « *« nécessaire, faisable et légitime le recours à de nouvelles compétences requises pour bien gouverner ou bien administrer. »*¹⁴⁶

Les politiques éducatives passent progressivement au cœur du débat public en France à la fin des années 1990, à la fois sous l'influence du Nouveau Management Public et sous celle grandissante du pouvoir médiatique sur la prise de décision politique. Ce « deuxième âge de la démocratie » selon la formule de Nathalie Mons¹⁴⁷ rend le citoyen, l'utilisateur, le parent, l'élève, légitimateur des décisions politiques. Celles-ci sont rapportées par les médias à grand renfort de communication de la part des acteurs politiques et servent de source pour mesurer l'adéquation des moyens employés par l'action publique avec les résultats obtenus et les objectifs qu'ils poursuivent. L'action publique rend des comptes au citoyen et est discutée publiquement. Les évaluations viennent soutenir scientifiquement les actions entreprises par les décideurs politiques et donnent un fondement rationnel jugé difficile à remettre en cause aux arguments invoqués dans les débats publics pour conclure à l'efficacité et à la pertinence de telle ou telle politique éducative. Les données statistiques deviennent les nouveaux outils favoris des décideurs pour étayer leurs conclusions et justifier les actions menées par le passé et celles à venir. Au sein du ministère de l'Éducation nationale, la Direction de l'évaluation, de la prospective et de la performance est créée pour contribuer en interne à l'évaluation statistique des résultats des élèves français.

¹⁴⁴ Charte de la société française de l'évaluation après les décrets de 1990 et 1998.

¹⁴⁵ Philippe Bezes, « Aux origines des politiques de réforme administrative sous la V^e République : la construction du « souci de soi de l'État » », *Revue Française d'administration publique*, 2, n°102, 2002.

¹⁴⁶ Philippe Bezes, *art. cit.*

¹⁴⁷ Nathalie Mons, « Evaluation des politiques éducatives et comparaisons internationales », *Revue française de pédagogie*, 164 juillet septembre 2008, p. 5.

Elle se voit à la fois attribuer le rôle de responsable des évaluations demandées par les ministres successifs et la possibilité de choisir de façon indépendante son programme de travail. Ce double positionnement, à la fois en lien avec les directives ministérielles et en autonomie pour ce qui est du choix des objets d'évaluation, n'a pas été sujet à questionnement dans les premières années d'existence de la DEPP. La proximité de conception de l'évaluation de l'action publique en éducation entre les directeurs de la DEPP et les ministres de l'éducation a permis, jusqu'à la fin des années 1990, de développer l'activité de la DEPP sans contestation franche de la validité des travaux qu'elle produisait alors et de leur légitimité à contribuer à orienter les politiques éducatives. L'arrivée de Claude Allègre au ministère constitue un véritable changement de perception de l'indépendance, et donc de la robustesse, des évaluations menées par des services statistiques internes. Le statut de direction ministérielle interroge le ministre sur l'indépendance des conclusions tirées dans les évaluations qu'elle mène, et son rôle se trouve largement réduit jusqu'au milieu des années 2000. L'apport d'expertise pour le pilotage des politiques publiques que constituent les évaluations des résultats des élèves est néanmoins crucial pour les décideurs politiques. La DEPP n'est pas seule à remplir cette mission, et ses agents doivent composer avec les interventions plus ou moins institutionnalisées d'acteurs aux pratiques et aux modes de fonctionnement divers.

L'arrivée sur la scène des évaluations des politiques éducatives de nouvelles évaluations « benchmark » internationales dans les années 1990 apporte une nouvelle dimension à l'évaluation de l'action éducative. Elle est désormais, pour les observateurs et usagers, comparable avec celle menée chez nos voisins plus ou moins proches – structurellement et géographiquement. Alors que les corps évaluateurs en France étaient traditionnellement des corps d'État (corps d'inspection, divisions ministérielles ou organismes publics chargés des traitements des données statistiques, Cour des Comptes etc.), la fin du XX^e siècle voit émerger de nouvelles institutions évaluatrices, y compris dans le champ des politiques sociales, culturelles, et éducatives : les organisations internationales. L'internationalisation des politiques publiques a redistribué les cartes : de

nouveaux acteurs sont entrés pleinement dans le jeu du pilotage des politiques publiques y compris éducatives, et se sont constitués en objet d'étude pour comprendre la formation des politiques publiques et l'évaluation de l'action publique.

Chapitre 2 : La création de PISA : « l’œil du Persan » ou l’externalisation de l’expertise des politiques éducatives

C’est donc dans un cadre évaluatif récemment institué en France que la fameuse étude PISA apparaît à la fin des années 1990. Elle est développée par l’OCDE, une organisation internationale économique à la légitimité forte dans son champ de compétence, mais qui n’avait jusqu’alors que peu investi le champ de l’éducation. En outre, lorsqu’elle l’avait investi, elle avait adopté un rôle de rassembleur de données au niveau des systèmes, en laissant les évaluations des élèves au soin des ministères de l’éducation nationaux. Et voilà que PISA est lancé, à contre-pied de tout ce qui avait été fait jusqu’alors au sein de l’organisation, et à contre-pied de la fonction même de l’OCDE. Elle qui n’a aucun pouvoir coercitif, qui ne peut pas légiférer, qui regroupe des pays aux organisations politiques, administratives et sociales très diverses, travaille avec PISA à produire des évaluations qui sont traditionnellement utilisées par les pays pour rendre des comptes à leurs administrés sur la façon dont est conduite et construite leur politique éducative. Malgré ce décalage de compétences, dont on aurait pu attendre qu’il engendre la méfiance des pays pour cette étude qui remettait en cause les compétences de chacun, l’évaluation PISA est devenue véritablement incontournable vingt ans plus tard. Et ce aussi bien pour les pays de l’OCDE que pour tous les autres, dont beaucoup cherchent à participer même s’ils ne sont pas membres de l’organisation¹⁴⁸.

Afin de comprendre comment l’OCDE a pu prendre une place de choix au sein des évaluateurs des politiques éducatives à la fois au niveau international mais aussi au sein des systèmes d’évaluation nationaux, nous observerons

¹⁴⁸ En 2020, l’OCDE compte 37 pays membres, et 79 pays ont participé à PISA en 2018.

d'abord le processus de construction de cette légitimité dans un secteur qui n'était pourtant pas le sien. Nous verrons dans un premier temps comment l'organisation économique est progressivement entrée dans le champ récent de l'évaluation des politiques éducatives, en prenant appui sur son rôle historique de coordinateur de la reconstruction économique des pays européens après la seconde guerre mondiale, puis en teintant des couleurs de sa doctrine économique sa conception des objectifs de l'éducation. Ce nouveau rôle qui se construit amène l'organisation à interagir avec les évaluateurs nationaux « historiques » d'une façon bien particulière. En se posant en répartiteurs des expertises de chacun tout en endossant elle-même une posture d'expert de l'évaluation des politiques éducatives, elle rebat les cartes du rôle traditionnel de l'organisation internationale sans pouvoir coercitif.

Nous regarderons ensuite dans un second temps plus attentivement l'objet « PISA » pour en définir les contours méthodologiques et les principes intrinsèques. L'OCDE se pose avec PISA en « œil du persan », en évaluation externe complémentaire des études nationales, qui mettrait en lumière, grâce à une neutralité statistique, les résultats de l'éducation des pays qui y participent. Charge à chacun ensuite d'en tirer ses conclusions en termes d'implication de politique éducative, du moins en théorie. Cette distance revendiquée de toute subjectivité nationale est notamment assurée par la mise en place d'un filet méthodologique complexe faisant intervenir les experts de l'évaluation nationale des pays participants à de multiples moments et à des multiples niveaux de responsabilité.

Le processus de création et les méthodes choisies pour PISA en font un objet nouveau dans le panel d'évaluations des politiques éducatives existantes à la fin du XX^e siècle. Plusieurs concepts clé en émanent, et en particulier celui de la « performance » appliquée aux résultats des élèves. En liant économie, développement, et éducation, en définissant de façon nouvelle les objectifs de l'éducation et en portant sa vision auprès de ses pays membres, en introduisant le concept de performance et de compétence plus que d'acquisition des connaissances dans sa façon de concevoir l'objet de son évaluation PISA,

l'OCDE prend une position originale dans le champ du pouvoir scolaire et dans le champ de l'évaluation de la politique éducative française. La description de ce qui fait de PISA un outil d'un genre nouveau d'aide à la décision publique est donc un préalable à l'étude de la réorganisation des relations entre les évaluateurs nationaux et leurs nouveaux homologues internationaux.

1) L'entrée de l'OCDE dans le théâtre de l'internationalisation de l'expertise en politique éducative

Evaluation phare de ces deux dernières décennies, le Programme International pour le Suivi des Acquis des élèves, plus connu sous le nom de « PISA », est mené par l'OCDE et fait l'objet d'un cycle de tests tous les trois ans. Son fonctionnement standardisé se rapproche des évaluations des élèves menées à partir des années 1980 au niveau national par les pays pour accorder les certifications ou rendre des comptes aux citoyens et à la représentation nationale sur les résultats de leurs choix de pilotage de la politique éducative.

Avant de s'intéresser au fonctionnement de PISA à proprement parler et à la méthode employée par l'OCDE pour occuper une place de choix dans les évaluations des politiques éducatives, on peut s'interroger sur la façon dont une organisation aux actions sur le champ de l'économie est entrée dans le champ de l'éducation, qui est, comme on l'a vu, une prérogative nationale majeure. Cette question se pose d'autant plus que PISA n'est pas une étude qui reprend des indicateurs produits par les pays pour dresser des portraits statistiques du fonctionnement des différents systèmes, mais bien une évaluation des résultats des élèves de plusieurs dizaines de pays, et autant de systèmes éducatifs et de programmes scolaires. Le chemin pour arriver jusque-là a été très progressif.

a) Un acteur original au sein d'une « communauté globale » étudiant les politiques éducatives

Le développement de l'OCDE s'inscrit dans le contexte plus large de la constitution d'une « communauté globale » selon la formule d'Akira Iriye¹⁴⁹, au sein de laquelle les organisations internationales occupent un rôle central d'« *espaces sociaux et culturels propices à la circulation d'informations, d'expertises et de savoir-faire.* »¹⁵⁰ L'après-Seconde guerre mondiale a ainsi été le terreau fertile du développement d'une nouvelle scène internationale d'organisations intergouvernementales travaillant sur un panel de thématiques aux implications politiques, et prenant la forme d'instances dotées de pouvoirs plus ou moins contraignants. L'OCDE en est l'une d'elle, et le rôle qu'elle acquiert au cours de ses premières années d'existence dépasse celle qui lui a été donnée au moment de sa naissance. Organisation internationale créée à l'après-guerre pour déployer en Europe les mesures du Plan Marshall, l'OECE devient OCDE en 1960. Elle ne dispose pas de pouvoir coercitif mais se pose en acteur du consensus international sur des sujets principalement économiques.

La trajectoire idéologique du travail de l'OCDE sur les questions de politique éducative est intéressante et permet de dégager de premières pistes d'analyses pour comprendre l'organisation du théâtre dans lequel s'affrontent les évaluateurs nationaux français et les « experts » de l'OCDE. Organisation économique, ses principes d'action ont visé dès le départ à favoriser la croissance des économies et par là leur prospérité, en créant un espace de coopération entre les gouvernements des pays membres – tous européens mis à part deux pays d'Amérique du Nord, les Etats-Unis d'Amérique et le Canada¹⁵¹. Signal politique

¹⁴⁹ Akira Iriye, *Global Community: the Role of International Organizations in the Making of the Contemporary World*, University of California Press, Berkeley, 2002.

¹⁵⁰ Sandrine Knott, « Dynamiques de l'internationalisation : l'Allemagne et l'Organisation internationale du travail (1919-1940) », *Critique internationale*, 52, 2011, p.70.

¹⁵¹ Les signataires de la Convention de Paris sont la République Fédérale d'Allemagne, la République d'Autriche, le Royaume de Belgique, le Canada, le Royaume de Danemark, l'Espagne, les États-Unis d'Amérique, la République Française, le Royaume de Grèce, l'Irlande, la République d'Islande, la République Italienne, le Grand-Duché de Luxembourg, le Royaume de Norvège, le Royaume des Pays-Bas, la République Portugaise, le Royaume-Uni de Grande-

de la volonté de créer une organisation de consensus, chaque membre dispose, au sein du Conseil des pays membres, d'une voix. L'accord mutuel de tous les membres est requis pour que les décisions soient prises et les recommandations validées. S'il est précisé que les signataires œuvrent conjointement pour « *développer au maximum et utiliser plus efficacement leurs capacités et leurs possibilités pour réaliser la plus forte expansion possible de leur économie et améliorer le bien-être économique et social de leurs peuples* »¹⁵², le lien avec les politiques éducatives n'apparaît pas à ce stade de la formalisation des objectifs de l'OCDE. Ils sont tous orientés vers les politiques économiques et financières, comme le détail d'article 1 de la Convention :

« Article 1

L'Organisation de Coopération et de Développement Économiques (appelée ci-dessous l'"Organisation") a pour objectif de promouvoir des politiques visant :

- a) à réaliser la plus forte expansion possible de l'économie et de l'emploi et une progression du niveau de vie dans les pays Membres, tout en maintenant la stabilité financière, et à contribuer ainsi au développement de l'économie mondiale ;
- b) à contribuer à une saine expansion économique dans les pays Membres, ainsi que non membres, en voie de développement économique ;
- c) à contribuer à l'expansion du commerce mondial sur une base multilatérale et non discriminatoire conformément aux obligations internationales. »

Tout juste la formation professionnelle est-elle citée dans les moyens à disposition des signataires pour atteindre les objectifs fixés, au milieu du développement des ressources économiques, scientifiques, technologiques, des politiques financières et commerciales, du soutien économique et technique aux pays non membres¹⁵³.

D'autres organisations sont à la même période davantage identifiées comme travaillant sur la question de la coopération internationale en matière de

Bretagne et d'Irlande du Nord, le Royaume de Suède, la Confédération Suisse et la République de Turquie.

¹⁵² Convention relative à l'Organisation de Coopération et de Développement Economiques, faite à Paris le 14 décembre 1960.

¹⁵³ Article 2 de la Convention relative à l'Organisation de Coopération et de Développement Economiques.

politique éducative, avec un angle d'approche différent de l'angle économique. Si toutes, au sortir de la guerre, se donnent pour objectif de garantir la paix, les réflexions ne portent pas sur les mêmes pistes d'action. L'éducation, et notamment l'éducation obligatoire, reste plus volontiers la prérogative des organisations spécialisées, comme l'Association internationale pour l'évaluation des rendements scolaires (IEA) sur laquelle nous reviendrons, ou d'organisations qui travaillent sur le champ de la culture et de l'égalité des droits – parmi lesquels l'accès à l'éducation – comme l'UNESCO. A partir des années 1950, les politiques éducatives deviennent un objet de travail privilégié pour plusieurs organisations internationales qui adoptent une démarche résolument comparatiste¹⁵⁴.

L'internationalisation des problématiques éducatives par l'adoption d'un prisme comparatiste s'est justifiée d'abord par la nécessité de promouvoir la paix entre les Etats et de lutter contre la pauvreté. Historiquement, dès les années 1920, la Société des Nations tentait déjà d'amener les questions éducatives à un niveau international (et notamment franco-allemand) en instituant de nombreux bureaux et autres fédérations caractérisées par un activisme dense autour de la promotion de la paix et de la collaboration pédagogique¹⁵⁵. On peut citer parmi eux par exemple le Bureau International de l'Enseignement Secondaire créé en 1919, la *World Federation of Education Associations* créée en 1923 ou encore le Bureau international d'éducation créé en 1925. Ces bureaux ont trouvé une pérennité après la Seconde guerre mondiale en étant intégrés pour certains au sein de l'UNESCO, en particulier le Bureau international d'éducation, toujours actif, et devenu organe de l'UNESCO en 1969 pour travailler en qualité de centre international d'éducation comparée.

L'UNESCO naît pendant la Seconde guerre mondiale en 1942 de l'esprit des gouvernements des pays européens réunis en Conférence des ministres alliés de l'éducation. Elle devient véritablement dans la seconde moitié du siècle

¹⁵⁴ Marcel Crahay et Arlette Delhaxhe, « L'analyse comparée des systèmes éducatifs : entre universalisme et particularisme culturel », *Raisons éducatives*, De Boeck Supérieur, 2004.

¹⁵⁵ Damiano Matasci, « L'éducation, terrain d'action internationale : le Bureau international de l'enseignement technique dans les années 1930 », *Relations internationales*, 3, n° 151, 2012.

l'organisation de référence sur les sujets éducatifs qu'elle place au cœur de son identité. Elle décrit ainsi ses missions relatives aux sujets éducatifs :

« L'éducation transforme les vie et elle est au cœur de la mission de l'UNESCO qui est de bâtir la paix, éradiquer la pauvreté et promouvoir le développement durable. »

La constitution de l'UNESCO précise :

« Les guerres prenant naissance dans l'esprit des hommes et des femmes, c'est dans l'esprit des hommes et des femmes que doivent être élevées les défenses de la paix ».

Pour cette organisation, la coopération mondiale notamment sur les sujets éducatifs est encore une fois, comme pour la SDN, un moyen privilégié de prévenir le déclenchement de conflits d'ampleur.

Au moment de la constitution de l'OCDE, bien que l'UNESCO semble constituer l'organisation internationale de référence sur le sujet, disposant d'équipes touffues et d'un Institut de statistique propre, ses productions ne font que rarement l'objet d'un usage de la part du gouvernement français. La France adopte certes les grands principes définis par l'UNESCO, comme les Objectifs de Développement Durable (ODD) et notamment l'ODD4 « *Assurer l'accès de tous à une éducation de qualité, sur un pied d'égalité, et promouvoir les possibilités d'apprentissage tout au long de la vie* », mais cela reste une déclaration d'intention. Et pour l'UNESCO, au-delà de ce travail intense de promotion d'objectifs communs ambitieux, il n'a jamais été question de produire des évaluations standardisées.

Peu de lien est donc établi directement entre économie et éducation au sein des organisations internationales travaillant sur le sujet. Un bureau toutefois s'était notablement penché dans les années 1930 sur la possibilité de résoudre les effets dévastateurs de la crise économique de 1929 sur la croissance mondiale en s'intéressant spécifiquement à l'enseignement technique. A l'époque il n'est pas encore question de travailler sur l'enseignement obligatoire dans son ensemble. La massification de l'accès à l'enseignement n'a pas encore eu lieu et il n'en est pas encore question. L'initiative de la création d'un bureau international de

l'enseignement technique (BIET) est franco-belge. L'objectif est de créer un espace d'échange international sur les questions d'enseignement technique, en charge « *de préparer les questions qui seraient traitées aux Congrès annuels de l'Enseignement technique, de fournir des renseignements aux institutions qui les demanderaient, d'étudier les principales questions à l'ordre du jour de l'enseignement technique* » selon les mots de l'un de ses fondateurs Edmond Labbé¹⁵⁶. Le BIET est conçu comme un outil d'aide à la décision dans le contexte post-crise de 1929, où le chômage est concomitant de la baisse de la croissance économique dans les pays occidentaux. La question du chômage touche alors principalement la main d'œuvre non formée et l'enseignement technique semble être le niveau de réponse adéquat à cette crise internationale qui implique, pour les fondateurs du BIET, une réponse internationale. Il est alors constitué sous la forme d'un groupement d'experts, principalement de hauts responsables des ministères de l'éducation de l'époque. Leurs compétences reconnues et leur bonne connaissance du monde administratif font d'eux, aux yeux des décideurs publics, des « experts internationaux ». Forts de leurs réseaux personnels influents, ils développent une proximité d'action avec le Bureau International du Travail – le seul alors capable de produire des normes internationales – et envisagent la question éducative globale comme étant étroitement liée à celle du travail. L'ambition est pour le bureau d'inscrire leur travail dans le sillage des valeurs de la SDN d'ériger la coopération internationale comme vecteur de paix et de progrès¹⁵⁷. Ce lien direct fait entre éducation et économie ne s'élargira pas à l'enseignement obligatoire et général après-guerre. L'OCDE aura alors l'espace pour entrer dans le champ éducatif lorsqu'il sera question de développer les systèmes massifs d'enseignement obligatoire dans la seconde partie du siècle dans ses pays membres. L'intérêt de ses derniers pour les indicateurs éducatifs de l'OCDE

¹⁵⁶ Edmond Labbé dans une lettre au directeur du BIT, Albert Thomas, cité par Damiano Matasci dans « L'éducation, terrain d'action internationale : le Bureau international de l'enseignement technique dans les années 1930 », *Relations internationales*, 3, n° 151, 2012, p.40.

¹⁵⁷ Damiano Matasci, *art. cit.*

évoluera on le verra au gré des besoins, notamment créés par les différentes crises économiques qui se succèderont sur la période.

b) Les débuts de l'OCDE et les politiques éducatives : la création du CERI

L'OCDE est donc un acteur à part de la communauté globale qui se crée autour des politiques économiques après-guerre. Cela ne veut pas pour autant dire que l'organisation ne travaille pas presque dès le départ sur le sujet des politiques éducatives, de l'enseignement pré-primaire à l'enseignement supérieur. Elle adopte un angle économique encore peu diffusé dans les organisations internationales qui s'intéressent au sujet éducation. Il faut attendre près de trois décennies pour que les publications qui s'y rapportent commencent à avoir autant d'écho auprès des médias et des décideurs publics qu'auprès des directions traitant des sujets économiques.

Lorsque PISA est lancé par l'OCDE dans la décennie 1990, l'organisation travaille sur le sujet de l'éducation depuis déjà plusieurs dizaines d'années, via notamment une instance décentralisée, le CERI, Centre pour la Recherche et l'Innovation dans l'Enseignement. Le CERI naît en 1968 et est placé sous le contrôle de l'OCDE. L'organisation travaillait jusqu'alors de façon plutôt confidentielle à la promotion de la coopération internationale en éducation pour favoriser l'allocation optimale des ressources¹⁵⁸, via notamment son comité de l'éducation qui avait succédé à un premier comité du personnel scientifique et technique. Le Centre s'inscrit dans la dynamique des années 1960 qui voient l'essor des réflexions sur l'innovation pédagogiques¹⁵⁹, dans le contexte de massification de l'enseignement dans de nombreux pays membres de l'OCDE. Il se concentre alors à lier recherche et innovation, qu'il définit comme « *toute tentative visant consciemment et délibérément à introduire dans le système d'enseignement*

¹⁵⁸ Pierre Laderrière, « Organismes d'enseignement et de recherche », *Revue française de pédagogie*, volume 50, 1980.

¹⁵⁹ Michèle Tournier, « La formation des enseignants par la recherche, en France : quelques points de repère (1966-1986) », dans *Recherche & Formation*, n°2, 1987.

un changement dans le but d'améliorer le système». L'objectif d'action sur la construction des politiques publiques en éducation transparaît dans cette définition, sans que soit encore très clair ce qui est entendu par « l'amélioration » du système.

Si l'on regarde les objectifs assignés au CERI à sa création, ou ses commanditaires et financeurs, la réponse n'est pas plus limpide mais des éléments de compréhension de la direction prise par les travaux du CERI commencent à se dégager. Du point de vue de sa gouvernance, le statut du CERI est ambigu. Contrairement au comité de l'éducation, il est au départ financé par des fonds extérieurs à ceux de l'OCDE, avant d'être finalement intégré en 1971 au fonctionnement de l'organisation. Il est toutefois dès le départ sous entier contrôle du secrétariat général de l'organisation et doit rendre des comptes à un comité directeur constitué de représentants de tous les pays membres, auquel il présente son programme annuel de travail pour adoption. Le travail du CERI s'est combiné à celui du comité de l'éducation interne à l'OCDE et les deux instances ont été dotées de missions complémentaires et successives dans le temps. Alors que le CERI est chargé d'identifier les barrières à cette fameuse « amélioration » des systèmes d'enseignement et des pratiques pédagogiques, en développant des expérimentations possibles pouvant être mises en place dans les pays membres, le comité de l'éducation intervient lorsque les pistes de travail sont déjà clairement formulées pour définir et coordonner les actions à mener.

Les thématiques de travail du CERI témoignent dès 1968 de l'angle d'approche adopté par l'OCDE en matière de politiques éducatives. L'éducation n'est pas à proprement parler, dans les discours et les rapports, associée au maintien de la paix. Un premier échelon intermédiaire est réellement identifié comme étant l'objectif principal de la recherche menée par le CERI : l'accompagnement de la massification de l'accès à l'enseignement pour permettre la croissance économique au cours des années 1970. Apparaît clairement alors le cadrage privilégié par l'OCDE pour aborder les politiques éducatives. L'éducation est un préalable indispensable au développement économique, au moment d'un ralentissement des croissances économiques et d'une montée du

chômage dans ses pays membres. En cela, l'organisation internationale rapproche l'éducation de ses thématiques classiques de travail, au premier rang desquelles le développement économique. Le programme de travail de la phase 2 d'existence du CERI (1972 à 1974) indique ainsi que les travaux du centre sont concentrés en trois grands sujets : les rapports entre l'éducation et la société, pour permettre de développer des stratégies innovantes permettant que la croissance économique soit aussi bien quantitative que qualitative – sans encore que cela ne soit défini clairement, si ce n'est de répondre aux « *nouveaux besoins sociaux* »¹⁶⁰ sans préciser de quoi il retourne – ; les processus d'enseignement et d'apprentissage par le développement et les échanges d'innovations ; le renforcement des dispositions nationales et internationales concernant l'innovation en matière d'enseignement.

Après 10 ans d'existence, le CERI conserve ces grands sujets de recherche qu'il adapte à la fois aux demandes des pays avec lesquels il travaille selon l'évolution de la conjoncture économique et sociale, et qu'il oriente en fonction des conclusions de ses propres travaux. De premières grandes problématiques émergent, et seront reprises par la suite dans les objectifs d'évaluation assignés au programme PISA notamment. Si certaines restent dans le champ de la gestion optimale des ressources, comme l'organisation et le financement de l'enseignement pour certaines catégories d'élèves ou certains niveaux d'enseignement, ou l'adéquation entre la structure des systèmes éducatifs et l'évolution anticipée des modes de travail et de vie, d'autres s'apparentent davantage à des sujets traités par la sociologie de l'éducation. Le CERI cherche ainsi à étudier notamment l'égalité éducative entre les différentes catégories sociales, sous l'angle de l'explication des bons ou de mauvais résultats scolaires et du lien possible avec l'origine socio-économique des élèves. Cette question bien connue des sociologues de l'éducation sera des années après au cœur de la communication de l'OCDE sur les résultats de PISA¹⁶¹.

¹⁶⁰ Pierre Laderrière, *art cit*, p. 104.

¹⁶¹ Voir notamment au chapitre 6 de cette thèse la façon dont les indicateurs de PISA sur le lien entre performance des élèves et origine socio-économique sera au cœur d'une appropriation par

Les travaux de l'OCDE sur les problématiques éducatives ne restent pas dans le cercle restreint de l'organisation. Les travaux du CERI et plus largement tout ce qui touche au sujet éducation fait l'objet dès les années 1970 d'un bulletin triannuel, « *l'innovation dans l'enseignement : nouvelles de l'OCDE Paris* », disponible gratuitement sur demande auprès de la Direction des Affaires sociales, de la main-d'œuvre et de l'éducation. Progressivement l'OCDE développe ses projets ayant traits à l'éducation et multiplie ses publications propres en en faisant des rendez-vous réguliers, et donc plus susceptibles d'être attendus.

c) L'installation d'un nouvel acteur incontournable des politiques éducatives

Finale­ment, l'OCDE, Organisation de Coopération et de Développement Economique, commence à investir le sujet de l'éducation à la toute fin de la décennie 1980. Elle s'emploie d'abord à renforcer ses indicateurs sur l'enseignement dans les pays de l'OCDE via notamment la création du projet INES, « indicateurs des systèmes d'enseignement » en 1988. Créé à la demande des gouvernements des pays membres, le projet INES avait à l'origine deux objectifs : leur permettre de comparer la « performance » de leur système éducatif, et de suivre l'évolution de son efficacité dans le temps. Pour répondre à cette demande, une publication annuelle est lancée au tout début des années 1990 qui permet la comparaison des données « système » des pays membres organisée selon différents indicateurs et différents chapitres. La publication est menée par le CERI et le premier *opus* sort en 1993. L'objectif n'est pas de regarder les effets de l'organisation de tel ou tel système éducatif, mais plutôt de dresser des portraits actualisés et problématisés des structures différentes des systèmes éducatifs des pays de l'OCDE. L'organisation travaille alors à concaténer et à mettre en lumière différents indicateurs sur l'état des moyens financiers, humains et sur l'organisation matérielle des systèmes éducatifs de ses pays membres à

le ministre français de l'Education nationale Vincent Peillon pour construire les conditions de légitimation de sa réforme pour la refondation de l'école.

partir des données fournies par les ministères de l'éducation des pays. Il s'agit en fait d'une photographie régulière. Dans le contexte des crises économiques des années 1990, les données de la publication s'inscrivent dans une logique comparatiste qui met en parallèle l'organisation des systèmes donnant ainsi des repères notamment financiers aux pays en recherche d'économie des deniers publics. L'éditorial de la publication sortie en 1993 indique bien l'importance du contexte économique mondial dans la décision de l'OCDE de travailler sur les politiques éducatives comme « *une des stratégies destinées à favoriser l'emploi et à réduire le chômage* » et cite le *Rapport intérimaire emploi-chômage* du Secrétaire général de 1993 :

« Seule une population active bien formée et très adaptable peut permettre l'ajustement au changement structurel et la mise à profit des possibilités d'emploi nées du progrès technologique. Pour y parvenir, il faudra dans bien des cas procéder à la révision, sans doute radicale, du traitement économique des ressources humaines et de l'enseignement. »¹⁶²

Regards sur l'éducation est un véritable « produit » phare de l'OCDE et est volontiers qualifiée de « meilleure vente » par le Secrétariat général, preuve de l'intérêt que lui portent à la fois les pays membres qui y sont représentés que les pays non-membres. Ce rôle d'inspirateur des politiques éducatives est un véritable objectif stratégique de l'OCDE qui permet à des pays non membres correspondant à certains critères et avec une contribution financière d'être progressivement intégrés à *Regards sur l'éducation*. Par ailleurs, les publications sont accessibles à tous et seront, dès que les moyens techniques le permettront au début des années 2000, disponibles gratuitement en ligne dans les deux langues de l'organisation, français et anglais. Bien que *Regards sur l'éducation* rencontre un succès certain à chaque sortie, l'éducation reste encore une des prérogatives de la très large Direction de l'éducation, de l'emploi, du travail et des affaires sociales, plus connue par son acronyme DEELSA, y compris lorsque l'évaluation PISA est conçue à partir de 1994.

¹⁶² OCDE, *Regards sur l'éducation*, Editions OCDE, 1993, p.9.

Finalement, la direction de l'éducation (EDU) est créée en 2002 par une division en deux des prérogatives de la DEELSA. Les directeurs qui se succéderont continueront à étendre le travail de l'OCDE dans le champ des évaluations des politiques éducatives en utilisant cette fois l'outil PISA. Barry McGaw, directeur adjoint de DEELSA en 2002, est le premier à prendre la tête d'EDU pendant quatre ans. En poste à l'OCDE, il milite pour que PISA soit une étude phare pour l'OCDE et serve de fondement aux travaux de l'organisation sur l'éducation, et notamment l'éducation obligatoire. Membre du Groupe directeur du programme DeSeCo (Définitions et sélection des compétences) pour l'OCDE aux côtés d'Andreas Schleicher – considéré on le verra comme le « père » de PISA – et de trois hauts fonctionnaires des institutions parrainant le projet¹⁶³, il contribue par ce biais à ce que les méthodes de PISA soient au cœur de la stratégie de constitution d'un programme cohérent d'évaluation des compétences clés adopté par tous les pays de l'OCDE. Les conclusions des travaux du DeSeCo indiquent ainsi que :

« Le Groupe directeur du programme DeSeCo invite l'OCDE :
(...) – A RECONNAITRE que le développement des évaluations futures doit reposer sur les expériences des études actuelles, en particulier le Programme international pour le suivi de l'acquis des élèves (PISA) et l'Enquête sur l'alphabétisation des adultes et les aptitudes à la survie (AAAS). »

Il sera un utilisateur convaincu de PISA après son passage à la tête de la direction et s'en inspirera largement lors de ses interventions portant sur les réformes éducatives en Australie notamment¹⁶⁴, prônant leur usage comme celui d'un livre ouvert sur le champ des possibles en la matière¹⁶⁵. Son adhésion aux principes méthodologiques de l'évaluation dérive de son parcours à la fois éducationnel et professionnel. Diplômé en psychométrie dont PISA reprend

¹⁶³ Heinz Gilomen pour l'Office Fédéral de la statistique de Suisse, Eugene Owen pour le National Center for Education Statistics des Etats-Unis, Scott Murray pour Statistique Canada.

¹⁶⁴ Par exemple la conférence donnée le 24 août 2010 sur la conduite des réformes en éducation pour la Banksia Association de l'Université de Murdoch.

¹⁶⁵ Barry McGaw, « The role of the OECD in international comparative studies of achievement », *Assessment in Education: Principles, Policy & Practice*, 15:3, 2008.

comme on le verra les principes méthodologiques, il fût, de 1984 à 1998, directeur exécutif de l'*Australian Council for Educational Research* (ACER). Autrement dit, l'organisme statistique chargé au départ de mener l'évaluation PISA pour le compte de l'OCDE.

Il laisse la place en 2006 à Barbara Ishinger, qui restera à la tête de EDU jusqu'en 2014. Docteure en littérature française diplômée de l'Université de Heidelberg, elle détient également une thèse d'Etat en littérature africaine. Passée par l'Unesco où elle fût directrice de la division de la coopération culturelle internationale, la promotion et l'enrichissement des identités culturelles, elle continue à porter PISA en renforçant la communication sur les implications de l'évaluation en termes de politique éducative. Au premier rang de ses messages forts, la nécessité de prendre du recul sur le fonctionnement de son système éducatif et d'utiliser tous les outils (et tout particulièrement les données PISA) pour réfléchir à l'éducation de demain, pour « apprendre un certain nombre de messages ». Les systèmes éducatifs sont à ses yeux rouillés, peu enclins à évoluer au contraire d'autres institutions notamment économiques, et PISA est un outil idéal pour avancer et entrer dans l'éducation du XXI^e siècle. Le parallèle entre PISA et le dynamisme permettant l'adaptabilité aux nouveaux enjeux de l'éducation est posé avec toutes les précautions nécessaires :

« Bien sûr vous ne pouvez pas transposer un bon exemple directement à un autre pays, parce que les cultures et les systèmes sont différents. Toutefois, il y a certainement un nombre de messages à apprendre. Comment cela se fait-il que tant d'institutions, d'entreprises aient changé au cours des dernières décennies, mais que les écoles n'aient pas changé ? Que s'est-il passé ? Si vous montrez à votre grand-père la salle de classe de votre enfant, il dira que cela ressemble à sa salle de classe des années en arrière. Quelque chose n'a pas eu lieu là, qui a eu lieu partout ailleurs. »¹⁶⁶

C'est également sous sa direction que PISA passe du niveau macro au niveau micro, avec le développement de « produits dérivés » de l'évaluation. Le programme *PISA for schools* est lancé en 2013. Ce projet vise à donner à chaque

¹⁶⁶ Barbara Ischinger, « What Do International Organizations have to do with Education ? », actes du colloque l'Ecole apprend elle le monde ? , 4 juillet 2014.

école la possibilité de passer de façon autonome un test PISA pour qu'elle puisse, selon la formule de Barbara Ischinger lors d'une interview donnée à l'AEF en 2012, « *apporter un éclairage sur l'environnement d'apprentissage au sein des établissements et en classe, les relations élèves-enseignant, et les attitudes et la motivation des élèves à l'égard de l'apprentissage* ».

Barbara Ischinger quitte l'OCDE en 2014 et est remplacée par Andreas Schleicher, vu pour beaucoup, à l'intérieur de l'OCDE comme à l'extérieur, comme le père de PISA. Formé en physique, en mathématiques et en statistiques, il est fin connaisseur des évaluations standardisées internationales. Après avoir travaillé des années à l'Association internationale pour l'évaluation du rendement scolaire (IEA) et occupé le poste de directeur de l'analyse, il entre à l'OCDE au début des années 1990 pour travailler au CERI. Il devient, 25 ans plus tard, directeur de la Direction de l'éducation, et occupe un poste considéré comme « clé » pour l'organisation, en étant proche du Secrétariat général. Au-delà de son travail de coordination des études de la direction de l'éducation, il a une véritable mission de conseil stratégique et de communication pour faire sortir les évaluations menées du simple giron des premiers concernés. L'enjeu de faire du travail de la direction un sujet global, à la fois au sens où il est connu de l'ensemble des pays du monde et à la fois au sens où tout le monde doit savoir ce sur quoi travaille EDU est devenu déterminant. Il est écrit en toutes lettres dans la biographie de Schleicher disponible sur le site de l'OCDE :

« En tant que membre clé de l'équipe de direction de l'Organisation, M. Schleicher appuie les orientations stratégiques du Secrétaire général visant à élaborer des analyses et des conseils propres à promouvoir la croissance économique et le progrès social. Il promeut les activités de la Direction de l'éducation et des compétences sur la scène mondiale et favorise la coopération tant à l'intérieur qu'à l'extérieur de l'OCDE. Outre les examens thématiques et nationaux, les travaux de la Direction incluent le Programme international pour le suivi des acquis des élèves (PISA), l'Évaluation de l'OCDE des compétences des adultes (PIAAC), l'Enquête internationale de l'OCDE sur l'enseignement et l'apprentissage (TALIS), et l'élaboration et l'analyse d'indicateurs mesurant la performance des systèmes d'éducation (INES) ».

PISA est bien cité en première place des travaux à promouvoir, par le directeur d'une direction désormais placée aux côtés du secrétariat général de l'organisation à une position stratégique, toujours au service de la croissance économique. Après deux décennies d'existence, l'OCDE entre ainsi pas à pas dans un champ encore relativement jeune au niveau international, en créant une évaluation d'ampleur dont elle teinte les principes des couleurs de sa doctrine économique. Quel est alors l'objectif de l'OCDE ? L'organisation s'intéresse-t-elle à l'évaluation des systèmes éducatifs par l'évaluation des résultats des élèves ? Évalue-t-elle par ce biais l'action des décideurs publics ? Ou celle des enseignants ? Ces questions sur l'objectif de ces évaluations standardisées sont centrales car elles permettent d'identifier la place que s'attribue l'OCDE au sein de l'ensemble des évaluateurs des politiques éducatives et la façon dont elle légitime sa position. L'histoire de PISA – car l'évaluation dispose bien, on le verra, de sa propre histoire – permet de mieux comprendre le rôle théorique et le rôle réel que se donne l'OCDE au sein des multiples acteurs impliqués dans la construction et la conduite du programme.

2) La genèse d'une évaluation qui se veut inédite

a) Une évaluation dotée d'un mythe créateur

Le projet PISA a été porté dès le départ par Andreas Schleicher qui intègre l'OCDE en 1994. Tous les acteurs rencontrés travaillant à PISA ou utilisant les données PISA, qu'ils soient français ou internationaux, le désignent volontiers comme en étant le père fondateur. Sa formation scientifique et les travaux qu'il avait menés à l'IEA orientent son travail lorsqu'il arrive à l'OCDE. Schleicher a en effet travaillé sur l'évaluation TIMMS au sein de l'IEA et souhaite aller plus loin. Convaincu de l'utilité et du potentiel des évaluations standardisées menées au niveau international, il se lance dans un nouveau projet à l'OCDE avec une marge de manœuvre et des moyens qu'il estime ne pas avoir eu auparavant.

L'histoire de PISA, telle qu'elle est racontée, commence avec lui. Schleicher lui-même fonde l'histoire de la création du concept de PISA autour d'un moment clé et de son parcours d'étude et professionnel :

« L'événement qui m'a mené à penser à ce projet a été une réunion avec les ministres de l'éducation qui étaient assis autour de la table comme cela, une vingtaine de ministres, et chacun d'entre eux disait: « dans mon pays, nous avons le meilleur système éducatif du monde », ou bien, « nous avons encore un petit problème, l'année dernière nous avons fait telle ou telle réforme pour résoudre ce problème ». Il n'y avait en fait aucun dialogue entre les pays. Et mon parcours est très différent de ce mode de fonctionnement. Je viens du monde de la science. J'ai étudié la physique. Et en physique nous savons comment pensent les autres, nous testons des idées, nous avons une communauté d'experts, nous avons un langage commun. Je trouvais qu'en éducation nous n'avions pas de compréhension commune, de concepts communs. Et cela m'a beaucoup surpris. J'ai alors dit que nous avons besoin de faire quelque chose, de créer un miroir dans lequel les pays pourraient d'observer eux-mêmes, à la lumière de ce qu'avaient fait les autres. Et en fait ça a marché comme ça. Et ce qui est très intéressant c'est qu'à cette époque il n'y avait pas de dialogue international. Désormais, lorsque les ministres de l'éducation se rencontrent, ils s'interrogent : « comment avez-vous mené cette réforme réussie dans votre pays ? Quels ont été les obstacles que vous avez rencontrés ? » Il y a beaucoup de bons échanges, et pas seulement au sujet de la construction des politiques publiques mais aussi à propos de leur mise en œuvre. Et je pense qu'en ce sens PISA a vraiment changé la nature de l'éducation. Il en a fait un phénomène global.»¹⁶⁷

Schleicher donne à l'évaluation un objectif nouveau qui fait de PISA un

¹⁶⁷ « *The event that led me to think of this project was a meeting with ministries of education who were sitting around the table like this, twenty something ministers, and each of them was saying : “in my country we have the best education system in the world”, and “we had a little problem left, last year we made a reform to tackle this problem”. There was actually no dialogue among countries. And my background is very different, I come from science. I studied physics. And in physics we know how other people think, we test an idea, we have a community of experts, we have a common language. I found that in education we didn't have a common understanding of this. A common concept. And I was also very surprised, and I said that we needed to do something to create a mirror in which countries can see themselves, in light of other countries achievement. And actually it has worked like this. And what is very interesting is that in those times, there was no international dialogue. Now ministers of education when they meet they ask : “how did you do this successful reform in your country? What were the obstacles?” There are a lot of good discussions, not only about the policy design but also on the implementation. And I think, in that sense, PISA has really changed the nature of education. It has made education a global phenomenon. »*, Entretien avec Andreas Schleicher, Directeur de la Direction de l'éducation, OCDE, 16 janvier 2017, traduit par nous.

nouvel outil de politique publique. Au-delà d'être une évaluation avec des résultats lisibles individuellement pays par pays ou les uns par rapport aux autres, il est d'abord décrit par son créateur comme un objet d'échange entre les décideurs publics nationaux qui a un double objectif. Il permet d'une part de créer ou de renforcer des relations entre les différents participants, et d'autre part de calibrer les orientations de politiques éducatives comme peuvent le faire toutes les autres évaluations internationales.

Ce « mythe créateur » donne le sens symbolique que devait avoir PISA à son origine pour celui qui en a été le porte-voix depuis le début: un langage universel et porteur de l'objectivité du langage scientifique. L'usage de la statistique pour construire les résultats PISA permettrait de disposer de données incontestables et utilisables par tous quel que soit le contexte dans lequel elles sont employées. Dans son usage par les décideurs publics, l'évaluation devrait servir à s'autoévaluer et contribuer ainsi à la réflexion dans le cadre de la politique éducative nationale.

Toutefois, si Schleicher a bien porté PISA dès le départ et en a été l'artisan pour l'OCDE, l'origine du projet semble en fait davantage issue de la demande expresse de certains pays membres – et notamment des Etats-Unis – confrontés à un besoin de travailler sur la formation de leur future force de travail et souhaitant disposer de données sur ce qu'ils appellent les « résultats » de l'éducation. Néanmoins, Schleicher tirera l'inspiration de sa formation scientifique pour ce qui est de la recherche de la création d'une évaluation « universelle », standardisée, à la méthodologie et aux items identiques dans tous les pays participants. Présent à toutes les étapes de la création et du lancement de PISA, il reste le chef d'orchestre de ce programme nouveau dont il oriente les choix techniques et stratégiques grâce notamment à la crédibilité dont il dispose pour avoir participé activement aux évaluations TIMMS à son poste précédent. PISA ressemble d'ailleurs beaucoup à TIMMS comme on le verra par la suite, même si l'OCDE réussit à introduire dans ce champ assez classique de l'évaluation en éducation des nouveautés qui en feront sa force. Autre force fondamentale, elle fait intervenir l'ensemble des participants à toutes les étapes de son cycle de vie.

b) Un outil éducatif à visée économique qui tire sa force de la recherche permanente du consensus

i. Une première phase de concertation pour construire le cadre complexe de cette évaluation nouvelle

Les documents issus des réunions de concertation préalables au premier test menées notamment dans les années 1997 et 1998 témoignent de la volonté de l'OCDE d'une part, et des pays membres d'autre part, de travailler sur la question de l'éducation « produit » permettant de former la future force de travail nationale et internationale. Forte du succès politique de *Regards sur l'éducation*, l'OCDE souhaite s'engager plus loin dans le conseil apporté aux pouvoirs publics en matière de politique éducative et compléter la mise en lumière d'indicateurs construits à partir de données existantes recueillies au niveau national par les pays faite par *Regards sur l'éducation*. L'objectif affiché par le secrétaire général Donald J. Johnston¹⁶⁸ est alors de renforcer la dimension analytique des productions de l'OCDE¹⁶⁹ en matière de politiques éducatives, en mettant en place des indicateurs ad hoc qui aillent au-delà du descriptif. Il précise :

« La nécessité d'établir des critères internationaux pour l'évaluation des systèmes d'enseignement se voit renforcer actuellement par une réorientation progressive de l'intérêt du public et des gouvernements d'un simple contrôle sur les ressources et le contenu de l'éducation vers une attention portée sur les résultats. »¹⁷⁰

En 1995 le « réseau A » d'INES, sous l'égide des Etats-Unis, commence à élaborer un cadre stratégique et méthodologique pour lancer un programme d'un

¹⁶⁸ Donald D. Johnston a été secrétaire général de l'OCDE de 1996 à 2006. Diplômé en droit de l'Université McGill à Montréal, il y a par la suite enseigné plus de 15 ans le droit fiscal avant de devenir Membre du Parlement dans les années 1980. Il a exercé plusieurs fonctions ministérielles : Président du Conseil du Trésor, ministre d'Etat chargé du développement économique et régional, ministre chargé des sciences et de la technologie, ministre de la Justice et Procureur général du Canada. En désaccord avec la ligne de son parti, il soutient dans les années 1980 l'Accord de libre-échange nord-américain, ce qui l'amènera à quitter la vie politique.

¹⁶⁹ DEELSA/ED/CERI/CD/INES/GA/A(2000)1/REV1

¹⁷⁰ DEELSA/ED/CERI/CD(97)6-FRE, p.2.

genre nouveau pour l'organisation : PISA, qui permettrait de produire des indicateurs comparables entre les pays portant sur les résultats de l'éducation. L'objectif est de mesurer auprès d'élèves de 15 ans comment ces derniers sont préparés à affronter les défis de leur vie future. Des tests comprenant des questions à choix multiples et des questions ouvertes sont administrés dans tous les pays de l'OCDE souhaitant participer. Ils sont passés auprès d'un échantillon d'élèves appartenant à un échantillon d'établissements représentatifs de l'organisation scolaire de chaque pays. L'évaluation est globale, car elle s'adresse certes aux élèves, mais aussi à leurs enseignants, et aux établissements qu'ils fréquentent, via des questionnaires dédiés qui permettent de recueillir une série de données contextuelles et plus qualitatives. Un cycle d'évaluation est mené tous les trois ans, avec à chaque fois une discipline majeure et deux disciplines mineures choisies à tour de rôle parmi la littérature, les mathématiques et les sciences. Nous reviendrons plus tard sur ce choix disciplinaire qui témoigne bien de la nouveauté de PISA dans l'environnement des évaluations des résultats des élèves.

Les données recueillies se doivent pour l'OCDE de répondre à une série d'impératifs de robustesse méthodologique et économique pour pouvoir servir d'outil au service de la prise de décision de politique éducative des pays participants mais aussi des autres : fiabilité, pertinence politique, efficacité, ponctualité et rentabilité¹⁷¹. Ils répondent à plusieurs grands objectifs dans la lignée des principes du New Public Management : s'assurer, par la comparaison de ses propres résultats, de l'efficacité de l'organisation de son système éducatif, contribuer à la « *transparence de l'enseignement ainsi qu'à l'élaboration de législations et de politiques, et permettre de cibler et de motiver les réformes de l'enseignement et l'amélioration des écoles.* »¹⁷² La communication de la construction de PISA autour de ces grands principes a pour rôle de s'assurer de la participation des pays – qui contribuent à étendre, par leur implication, l'influence potentielle de l'évaluation et qui apportent des financements conséquents – et en premier ceux dont les

¹⁷¹ DEELSA/ED/CERI/CD(97)6-FRE, p.2.

¹⁷² *Ibid.* p.6.

économies sont dans des situations difficiles et qui commencent à appliquer les principes du New Public Management.

Une stratégie de collecte des données pour PISA est donc mise en place autour de quatre séries d'indicateurs, qualifiés par l'OCDE de « catégories de produits ». La première est composée des outils descriptifs permettant aux décideurs politiques de saisir l'image générale de trois composantes bien distinctes : les connaissances, les compétences et les qualifications de leurs élèves. C'est là le début de la formalisation de l'originalité de PISA par rapport aux évaluations standardisées existantes jusqu'alors, qui ne s'étaient pas encore penchées réellement sur l'évaluation des compétences des élèves et restaient concentrées sur l'acquisition du contenu d'un programme. La seconde est composée d'indicateurs contextuels permettant d'éclairer les résultats des élèves et de mieux saisir les déterminants de politique publique qui peuvent avoir un effet sur eux. On voit déjà poindre la volonté de l'organisation de ne pas se cantonner à un simple compte-rendu de résultats par pays, mais bien de donner aux décideurs publics des orientations argumentées statistiquement pour activer des leviers certes éducatifs mais aussi sociaux et économiques ayant un impact possible sur les résultats de leur politique éducative. La troisième catégorie de produits est constituée d'indicateurs de tendance, permise par la répétition tous les trois ans de l'évaluation. L'OCDE s'éloigne ainsi de l'évaluation ponctuelle et s'assure de l'intérêt renouvelé des pays sur une période longue. Cela donne également le pouvoir de se détacher de l'écueil du classement et de permettre ce fameux effet « miroir » souhaité par Schleicher, où chacun peut s'observer et se voir évoluer de cycle en cycle. La quatrième catégorie de produits est constituée de toute une série d'informations générales sur la vie dans les établissements, le bien-être des élèves, des enseignants, et tout autre questionnaire jugé intéressant pour fournir « de nouvelles analyses plus ciblées de politiques ». Le descriptif des résultats est décidemment insuffisant et l'OCDE souhaite disposer d'une base de connaissances lui donnant un pouvoir d'analyse des politiques éducatives.

Cette stratégie de collecte de données ambitieuse, qui sert de cahier des

charges pour le prestataire statistique qui sera retenu après appel d'offre pour les premiers cycles de l'évaluation, est validée en mai 1997 par tous les pays participants à PISA. Plus que simplement validée, elle a en fait été co-construite au cours de réunions sur plusieurs jours organisées par l'OCDE dans différents pays souhaitant participer à l'évaluation, et pas seulement dans ses locaux parisiens. Cette délocalisation des réunions est un moyen de choix pour renforcer le sentiment d'implication des pays dans la construction de PISA. Elle sera d'ailleurs la règle pour toutes les réunions qui auront lieu par la suite, quels que soient les acteurs en présence (experts techniques, acteurs politiques, décisionnaires stratégiques, etc.). Les délégations des pays concernés ont alors la main sur l'organisation de la venue des représentants de tous les autres pays intéressés et des équipes de l'OCDE, tout en restant dans le format classique des réunions internationales. Ces déplacements, très appréciés des participants, relativisent, au moins en termes d'image, la mainmise de l'OCDE sur l'évaluation et créent un sentiment d'appartenance à une « communauté PISA » facilitant la recherche du consensus au moment des prises de décision stratégiques.

En plus de l'évaluation socle, pour laquelle les pays contribuent largement financièrement, le choix leur est laissé de prendre part à tel ou tel questionnaire ou à telle ou telle option qui permet de creuser des points généralement considérés comme prospectifs. Cette modularité des participations laisse la marge de manœuvre aux pays de participer à PISA tout en se retirant des indicateurs qu'ils estiment ne pas correspondre aux spécificités de leur système national. C'est également un moyen, pour les pays, de maîtriser les coûts de participation à l'évaluation qui constitue souvent une dépense bien plus importante que toute autre participation aux autres évaluations internationales. Chaque option étant payante, elle sert parfois de variable d'ajustement pour ne pas grever des budgets d'administrations déjà souvent en restrictions budgétaires.¹⁷³

¹⁷³ Entretien avec Bruno Trosseille, Directeur du bureau des évaluations des élèves, DEPP, 14 avril 2016.

« On a quand même un assèchement des ressources financières de la DEPP pour participer à toutes ces études, au détriment du financement des études nationales. Et quand on regarde, la cotisation annuelle pour PISA c'est environ 300 000 euros. Pour TIMMS c'est 90 000 euros, et PIRLS c'est 40 000 euros. Nous on a lancé CEDRE, ça coûte environ 300 000 euros par cycle, mais ça a vocation à baisser rapidement. Pour l'instant une part importante du coût est due aux prestataires qui doivent gérer la numérisation des tests. Dans le futur on fera l'étude sur ordi. (...)

Pour PISA c'est la DEPP qui d'habitude les options qui vont être retenues par la France, en fonction de leur intérêt et du coût total du projet ».

« Le regard international est important mais il entraîne un désengagement au niveau national, notamment financier et notamment sur les évaluations sur les programmes ».

Ainsi, la France ne participe généralement pas par exemple au questionnaire passé auprès des enseignants, la DEPP considérant que les réponses qu'il apporte ne permettent pas de donner les éléments d'analyse considérés les plus pertinents pour expliquer les résultats obtenus.

« On n'utilise pas le questionnaire enseignant, ce n'est pas ce questionnaire qui est pertinent pour expliquer les résultats donc on ne prend pas cette option dans PISA. Les questionnaires chefs d'établissement par contre permettent de situer l'élève dans un contexte. »¹⁷⁴

Ce refus d'option réaffirme symboliquement le fait que PISA n'évalue pas, pour la DEPP, les pratiques enseignantes, mais bien les résultats des élèves dans un contexte d'enseignement particulier. Le chef d'établissement est alors au départ considéré comme le mieux à même de fournir les éléments d'explication pertinents pour décrire ce contexte en donnant à voir l'organisation des enseignements de façon globale à l'échelle de l'établissement, sans voir la façon dont la pédagogie est pratiquée dans la classe.

¹⁷⁴ Entretien avec Bruno Trosseille, Directeur du bureau des évaluations des élèves, DEPP, 14 avril 2016.

C'est enfin une manière pour les pays de manifester leur désaccord avec des partis-pris méthodologiques ou d'analyse choisis par l'OCDE et rendus publics au moment de la publication des résultats sans que les pays n'aient forcément eu la possibilité d'avoir leur mot à dire sur la version finale de l'exploitation des données.

Ces participations à géométries variables se font au vu et su de tous les autres participants et ce pouvoir de positionnement, ou plutôt de non-positionnement, est donc utilisé avec grande parcimonie car il pourrait être perçu comme une volonté de masquer les résultats les moins favorables. En pratique cela a pu être le cas de la France qui a régulièrement retiré sa participation à certaines options voire retiré ses données une fois les indicateurs créés. Le choix de participer ou non est fait, en France, au niveau de la DEPP. Quelques ordres contraires ont pu être donnés par le cabinet du ministre lorsque les options choisies ou au contraire rejetées étaient porteuses de représentations des orientations politiques souhaitées pour les réformes éducatives à venir. Ce fût par exemple le choix, lors du cycle 2009, où l'OCDE proposait d'inclure une évaluation de la littératie financière des élèves. La DEPP choisit alors, compte tenu de l'absence totale de ce sujet dans les programmes enseignés aux élèves avant 15 ans, de ne pas retenir cette possibilité payante. Validée par le directeur de cabinet du ministre Luc Chatel, la décision est toutefois remise en question par un conseiller du ministre souhaitant travailler sur le sujet à l'avenir avec l'inspection générale d'économie et souhaitant se servir des résultats des élèves comme d'une base de réflexion sur les choix pédagogiques à faire.

Cette marge de manœuvre laissée théoriquement aux pays dans le choix des options auxquelles ils participent, compte tenu des implications fortes qu'elles peuvent avoir dans l'élaboration de la politique éducative des pays concernés, contribue à fabriquer la stratégie du consensus souhaitée par l'OCDE dans la mise en œuvre de l'outil PISA.

ii. Le lancement de PISA : une stratégie du consensus revendiquée

Au-delà du processus même de décision imaginé par l'OCDE, la stratégie du consensus menée par l'organisation s'est matérialisée par la mise en place d'un filet méthodologique fort, et de phases de test systématiques pour s'assurer de l'adhésion de tous les participants une fois l'évaluation lancée sur le terrain. Il s'agit d'anticiper autant que possible les critiques qui pourraient affaiblir la fiabilité des résultats et réduire leur utilisation par ceux qui en seraient devenus les détracteurs.

La première phase de test est organisée en amont du premier cycle de PISA, notamment pour s'assurer que la coordination entre les différents acteurs est suffisamment efficace pour pouvoir tenir les délais et le budget alloué à chaque cycle, et pour vérifier de la pertinence des résultats obtenus pour pouvoir servir d'étai aux décisions de politiques éducatives nationales¹⁷⁵. Ce fonctionnement « essai – évaluation » perdurera pour tous les cycles d'évaluation. L'objectif est avec ça d'éviter au maximum les biais liés à des particularismes locaux et de prendre en compte autant que possible les commentaires des pays une fois les retours sur les phases d'essai formulés¹⁷⁶. C'est un moyen pour l'OCDE de s'assurer de la validité statistique des résultats obtenus par les pays par le biais notamment de la validité de l'échantillon choisi, mais aussi de l'adhésion des pays au programme à travers la mobilisation suffisante des établissements pour permettre la représentativité des résultats et la reconnaissance de la validité des contenus des tests passés.

Ces mécanismes d'évaluation sont coordonnés par le Groupe de direction du projet INES dès le départ et selon un processus cadré. Celui-ci demande en effet dès février 1999 au Secrétariat général de constituer un plan d'évaluation de PISA avec le CPP et les instances chargées, pour l'OCDE, de produire les travaux sur les acquis de l'enseignement. Une fois ce travail préalable réalisé, censé

¹⁷⁵ DEELSA/PISA/BPC(99)9.

¹⁷⁶Entretien avec Sophie Vayssettes, responsable de projet PISA, OCDE, 5 novembre 2015.

garantir la validité a priori de la méthodologie retenue, le premier cycle d'évaluation est mis en place au premier semestre 2001. Les résultats sont recueillis avant la fin du mois de juillet avant d'être traités et publiés sous la forme d'un rapport d'un volume modeste pour l'organisation, d'un peu moins de 200 pages hors annexes et description méthodologique. Ce premier volume reprend de façon relativement descriptive les résultats de l'évaluation. Il donne les performances en compréhension de l'écrit et en culture mathématique et scientifique des élèves des différents pays participants, décrit les résultats de l'évaluation des approches des élèves « *à l'égard de l'apprentissage et leur confiance en leurs propres capacités.* »¹⁷⁷ Au fur et à mesure les publications croîtront de façon exponentielle pour intégrer de plus en plus d'analyse problématisées. Le dernier cycle PISA 2018 fait ainsi l'objet de six volumes, dont un seul donnera les résultats des élèves aux tests. Les autres volumes sont constitués d'indicateurs produits par l'OCDE à partir des résultats, mettant en lumière des sujets de politique éducative comme les écarts de performance entre les garçons et les filles selon les disciplines, la capacité des systèmes éducatifs à réduire les inégalités sociales et à permettre aux enfants issus de l'immigration de suivre une scolarité avec succès, ou encore la façon dont les élèves se sentent à l'école et l'effet que le bien-être scolaire peut avoir sur les résultats en classe. L'OCDE prend progressivement sa place d'analyste experte des politiques éducatives tout en se positionnant en coordonnatrice de l'action d'une pluralité d'acteurs présentés par elle comme ceux « faisant » PISA.

c) Les acteurs de PISA : une expertise en mille-feuilles

La construction de PISA fait intervenir pléthore d'acteurs, et une gouvernance pluricéphale est mise en œuvre dès le départ. L'OCDE revendique sa volonté, comme pour ses autres travaux, de faire de PISA un programme coopératif au service des pays participants, dont elle assurerait la coordination et la publication

¹⁷⁷ OCDE, *PISA 2000*, Editions OCDE, 2001, p.106.

des résultats. Cette volonté s'inscrit dans la lignée de ce que le Secrétariat général avait entrepris depuis 1995 et la création du projet INES, voulant renforcer les indicateurs de l'enseignement produits par l'OCDE. Le Secrétaire général indiquera ainsi dans une note publiée plus tard en 2000 à l'occasion de la publication du programme de la quatrième assemblée générale du projet INES que « *l'amélioration des mécanismes de coordination et l'assouplissement du système de réseaux ont permis aux pays Membres de prendre part plus efficacement à la mise au point, la publication et à l'analyse des indicateurs de l'enseignement de l'OCDE.* »¹⁷⁸

Dans le cadre de PISA, une attention particulière est mise à assurer, par écrit, que parmi les objectifs centraux de l'évaluation se trouvent à la fois la représentativité des pays participants par l'organisation de réunion et de collaborations « amples », tout en maîtrisant les dépenses et les délais et en s'assurant que les intérêts de politique nationale de quelques-uns ne mettent pas en danger l'objectif de comparabilité internationale central dans l'évaluation¹⁷⁹.

Au sein du programme PISA, les pays sont finalement représentés et ont officiellement leur mot à dire sur les décisions prises – aussi bien politiques que techniques – à toutes les étapes de la construction de PISA, de l'organisation des cycles d'évaluation et de la mise en forme et de l'exploitation des résultats.

i. Le Conseil des pays participants, un acteur stratégique central

La tête politique de PISA est le Conseil des pays participants (CPP), ou *Pisa Governing Board* en anglais. Il est mis en place en 1997, au début du processus d'élaboration des principes directeurs de PISA, dont le premier vrai cycle d'évaluation – sans compter l'incontournable test préliminaire - est prévu en 2000. Le CPP est composé d'un expert en évaluation des connaissances choisi par chaque pays parmi ses experts nationaux. Son rôle est d'élaborer le programme de travail de PISA, qui est alors statutairement un programme

¹⁷⁸ DEELSA/ED/CERI/CD/INES/GA/A(2000)1/REV1, p. 2.

¹⁷⁹ DEELSA/PISA/BPC(99)9-FRE, p. 5.

décentralisé, et d'en suivre le bon déroulement. C'est lui qui décide de la stratégie à adopter pour élaborer de façon périodique les indicateurs sur les résultats des élèves. Car pour l'instant ce sont bien les « résultats » qui intéressent les pays, sans que leur nature ne soit réellement précisée. On ne parle pas encore de « performance », du moins le terme n'est pas encore employé communément dans les comptes-rendus des réunions de préparation ou les notes de travail. Le CPP n'intervient pas tout seul dans le processus de décision. Il rend compte de ses choix au Comité de l'éducation et au Comité directeur du CERI en passant par le Groupe de direction du projet INES.

Le CPP est institué en septembre 1997 et se réunit pour la première fois pour élire son président et les vice-présidents, désignés par les pays participants et élus pour trois ans. Il adopte à cette occasion les guides qui organisent son action.

Plusieurs questions stratégiques sont abordées dès le départ, notamment celle de la population cible des tests, paramètre éminemment politique puisqu'elle est le déterminant de ce qui pourra être attendu des élèves. Ce paramètre n'implique pas seulement de définir jusqu'à quel niveau de contenu scolaire les tests pourront aller, mais en creux il demande aux pays de définir le véritable objectif de l'évaluation internationale. Que veut-on évaluer communément, alors que tous les pays ont leurs propres programmes scolaires, parfois déterminés au niveau national, parfois régionalement, parfois laissés dans une certaine mesure à la discrétion des établissements scolaires eux-mêmes ? Cette question est véritablement au cœur de la construction de PISA et la réponse apportée en fait un objet nouveau dans le champ de l'évaluation. Le choix est fait de ne pas évaluer les résultats des élèves arrivés à un niveau scolaire donné, mais plutôt à un âge donné. Dès lors, le prisme méthodologique choisi est celui de se détacher de la notion de programme, qui ne permettrait pas de rendre les résultats comparables entre les différents systèmes. Les tests évalueront les « acquis » des élèves, sous l'angle de leur capacité à s'en servir au-delà des situations scolaires. L'OCDE s'écarte alors de la classification standard de l'acquisition des compétences mise au point entre 1965 et 1975 par l'UNESCO et le BIT, la

Classification Internationale Type de l'Éducation (CITE) qui instaure en unité de classement le programme éducatif¹⁸⁰.

Deux mois plus tard, le CPP se réunit à nouveau à Paris afin de mettre en place les différents groupes d'experts fonctionnels nommés pour travailler à la conception des tests et de leur cadre conceptuel. Trois grands domaines sont retenus, pour plusieurs raisons. Dans la mesure où les acquis doivent permettre aux élèves de devenir de futurs acteurs productifs de la société, et où ils doivent de surcroît pouvoir être évalués dans l'ensemble des pays participants dans que le contenu des programmes ne puisse interférer, le choix est fait de se concentrer sur les savoirs fondamentaux que sont la lecture et les mathématiques, et d'évaluer les sciences. La volonté est celle de s'inscrire dans une démarche de productivité prospective, et d'étudier la possibilité de parler un langage commun malgré les différences linguistiques. Les sciences deviennent donc le troisième pied d'une évaluation prévue en trois éditions à trois ans d'intervalle. Un quatrième groupe est institué, chargé de concevoir le questionnaire contextuel qui permettra de donner du sens statistique et explicatif aux résultats des élèves. L'objectif d'analyse fine des effets de l'éducation et d'outillage de la prise de décision implique pour l'OCDE de mettre en lumière différents paramètres pouvant influencer sur les résultats au-delà des données système (taille des classes, nombre d'heures d'enseignement, etc.). Le contexte socio-économique à la fois des élèves et des établissements retenus dans l'échantillon de l'évaluation acquiert ici une place centrale et des questionnaires ad hoc sont donc prévus.

ii. Des groupes experts passeurs de leurs paradigmes

Les groupes experts sont constitués rapidement, dès la fin de l'année 1997, et comportent entre 8 membres (pour ceux de sciences et de mathématiques) et

¹⁸⁰ Annick Kieffer et Rémi Tréhin-Lalanne, « La fabrication d'un consensus : la révision de la Classification Internationale Type de l'Éducation », *Sociologie et sociétés*, Vol.43, 2, automne 2011.

10 membres (pour ceux de lecture et du questionnaire contextuel). Ils passeront progressivement à une douzaine de membres au fur et à mesure des cycles, avec un turn-over important. Ils ont un rôle tout à fait central et leur choix est hautement politique car s'ils ont pour mission de construire un cadre d'évaluation applicable dans tous les pays participants sans n'en favoriser aucun, ils sont teintés particulièrement des principes des systèmes éducatifs de leurs pays respectifs et de leurs propres travaux de recherche. Chaque groupe est mené par un responsable issu d'instituts d'évaluations ou d'universités. En pratique, tous les pays participants ne sont pas « représentés » par ces experts et il est difficile de savoir comment ils sont choisis. Les documents d'archive de l'OCDE consultables ne le précisent pas. C'est pourtant un choix porteur de sens car les experts choisis ont apportés avec eux leurs paradigmes et ont parfois insufflé aux travaux préliminaires sur PISA l'esprit et les méthodes des études auxquelles ils avaient auparavant contribué, et qui avaient été parfois fortement remises en question par certains des pays participant à PISA. C'est particulièrement le cas du premier groupe d'experts sur la lecture dont le responsable (le « chairman » en anglais) est le Professeur Irwin S. Kirsch, membre de *l'International Testing Service*. Il a travaillé dès les années 1980 au développement d'évaluation sur l'acquisition de la lecture par les adultes notamment, y compris en prenant la tête d'une évaluation réalisée par l'OCDE et dont les résultats ont été publiés au début des années 1990 : IALS, pour « International Adult Literacy Survey ».

Ce point de parcours est intéressant à observer car il est porteur de sens pour comprendre les différences d'orientation, notamment méthodologiques, qui ont pu surgir dès le départ entre l'évaluation des résultats menée jusque-là en France et l'optique voulue par l'OCDE. La France avait participé à l'évaluation, jugée alors par la DEP comme assez nouvelle, traitant d'un sujet récemment apparu sur le devant de la scène des politiques publiques d'éducation et de formation. Jamais réellement évaluée jusque-là au niveau national, les compétences des adultes en littératie, concept dont la définition sera formalisée par l'OCDE en 2000 comme « *l'aptitude à comprendre et à utiliser l'information écrite dans la vie courante, au travail et dans la collectivité en vue d'atteindre des buts personnels*

*et d'étendre ses connaissances et ses capacités*¹⁸¹», suscitait l'intérêt du ministère et du directeur de la DEP à l'époque, Claude Thélot. Sa direction n'a cependant pas été chargée de suivre l'étude, qui a été déléguée au secteur privé car l'âge adulte ne relevait pas directement de l'éducation nationale.

Les documents de travail sont toutefois transmis à la direction qui les étudie pour formuler un avis sur le déroulement de l'évaluation, avec l'aide de l'INED. Dès le départ, la question de la façon d'évaluer s'est posée pour la DEP : pas question d'évaluer les adultes, généralement dans le contexte du domicile, de la même façon que l'on pouvait évaluer des élèves dans un contexte scolaire de salle de classe. Les appréhensions méthodologiques de la direction n'ont pas engendré de modification spécifique de la façon dont l'évaluation a été menée par l'OCDE. Si elle n'a officiellement pas eu son mot à dire puisqu'elle n'était pas saisie dans le cadre de IALS, difficile de savoir si elle n'a pas été entendue ou si elle a réagi trop tardivement. Toujours est-il que ses craintes se sont matérialisées concrètement au moment du rendu du rapport : les résultats de la France à IALS ont été catastrophiques. Selon l'évaluation, près de 40% des adultes de 18 à 65 ans en France étaient en grande difficulté devant l'écrit en 1994. Face à ces résultats, qui avaient fuité dans la presse avant la publication officielle du rapport, la France a décidé à l'époque de se retirer de l'évaluation. Pour la DEPP qui revient quelques années plus tard sur le sujet, les résultats sont liés de façon certaine à une mauvaise méthodologie de recueil des données. Ils indiquent dans une publication sortie en 2008 et portant sur les méthodes d'évaluation des compétences des adultes : *« des expertises méthodologiques ont remis en cause la fiabilité de l'enquête, justifiant le retrait de la France de l'opération. Cet article reprend l'un de ces travaux, en montrant l'importance des conditions de collecte. (...) Les instances concernées par le sujet ont décidé en 2000 de privilégier une approche nationale, pour développer nos connaissances dans ce domaine très particulier »*. C'est la naissance de l'enquête Information et Vie Quotidienne (IVQ), menée par l'INSEE et pilotée au départ conjointement avec un ensemble d'organismes, l'ANLCI, Agence nationale de

¹⁸¹ OCDE, *La littératie à l'ère de l'information*. Rapport final de l'Enquête internationale sur la littératie des adultes, Editions OCDE, 2000.

lutte contre l'illettrisme, le CGP, Commissariat général au plan, le CREST, Centre de Recherche en Economie et Statistique, la Dares pour le ministère du Travail, et l'INETOP, l'Institut national d'étude du travail et de l'orientation professionnelle.

La France va plus loin dans le désaveu des études internationales menée par l'OCDE et décide quelques années plus tard de ne pas participer à celle prévue par la suite, portant sur le même sujet, et intitulée ALL, Enquête sur la littératie et les compétences des adultes. Une expertise est menée en interne de la DEP en amont et sert de fondement à ce refus. La réponse est claire, la France n'a pas le contrôle suffisant sur cette évaluation dont la méthodologie est toujours discutable. Un cadre de la DEPP, chargé à l'époque de mener ces expertises, précise ainsi que : « *l'expertise que j'avais faite c'est qu'il y avait plein de problèmes sur ce genre d'enquêtes, et en 1998 le problème n'était pas résolu. Le conseil c'était alors de ne pas s'inscrire, on ne savait pas comment allaient être les résultats.*¹⁸² »

Ce choix politique a des conséquences sur l'ampleur prise par cette nouvelle évaluation. Son retrait n'inspire pas forcément confiance aux autres pays membres de l'organisation et est même, pour certains cadres de la DEPP, la raison pour laquelle l'évaluation n'a pas rencontré le succès attendu. La France aurait à ce moment retrouvé son moyen de pression sur l'organisation internationale pour plusieurs raisons, liée à sa localisation géographique et au statut de puissance économique de la France.

« L'OCDE est en France, c'est donc une grosse pression. On est un des gros pays de l'OCDE. Qu'une des cinq premières puissances ne veuille pas participer ça interroge pour recruter les autres. ALL a très peu marché, avec 12 pays inscrits dont par exemple les Bahamas, des tous petits pays. A partir de là, la France a commencé à développer ses propres enquêtes, IVQ par exemple. On fait des critiques, ils ne bougent pas, donc on fait nos enquêtes, qui ont été un grand succès. »¹⁸³

On voit bien que de vrais rapports de force se dessinent, cristallisés en

¹⁸² Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 21 mars 2019.

¹⁸³ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 21 mars 2019.

surface autour de la méthodologie d'étude. L'OCDE d'ailleurs travaille le sujet avec la plus grande attention que ce soit en amont de la première évaluation ou à chaque cycle, où elle s'efforce de détailler la méthodologie employée dans tous ses supports de communication. Des chapitres entiers, puis des volumes au fur et à mesure que les cycles se succèdent, y sont d'ailleurs consacrés. Mais finalement, au-delà de la méthode, il transparaît dans les choix qui sont faits par la France que ce sont bien les influences sur les évaluations qui sont au cœur du conflit. Le rôle de son « National Project Manager », qui la représente à la fois auprès de l'OCDE et auprès de tous les autres pays participants à l'évaluation, est alors primordial.

iii. Les National Project Manager, un maillon central pour les intérêts nationaux

Un rouage essentiel est institué pour assurer la participation à la fois politique et technique des pays : le chef de projet national, plus connu sous l'acronyme « NPM » pour *National Project Manager*. Représentants de leur administration nationale chargée d'administrer les tests, ils sont choisis par les responsables des évaluations en éducation de chaque pays participants. Dans le cas de la France, la DEPP hérite de la mission de suivre l'évaluation PISA et d'organiser la passation des tests dans les établissements scolaires sélectionnés dans l'échantillon français. Les NPM français sont issus de ses rangs. Plusieurs se sont succédés au fil des évaluations, selon les disciplines fortes de chaque cycle. Il s'agissait à chaque fois de trouver des interlocuteurs non seulement compétents en statistique appliquée à l'évaluation des politiques publiques, mais aussi capables de fournir une opinion sur le contenu disciplinaire des tests. Ils ne sont pas situés à un échelon décisionnel au sein de la DEPP mais travaillent concrètement au bon déroulement des évaluations nationales et internationales dans lesquelles la France est par ailleurs engagée.

Par le rôle que l'OCDE et le consortium responsable de l'organisation technique des épreuves PISA leur assignent, les NPM se trouvent à l'interface

entre l'organisation et leur administration centrale. Porte-parole bidirectionnels, ils font le lien entre les propositions de l'OCDE et les orientations politiques souhaitées au niveau national. S'ils ne prennent pas de décisions stratégiques, ils sont les ambassadeurs techniques de leur ministère à toutes les réunions organisées par l'OCDE et échangent entre eux sans forcément toujours attendre les rendez-vous officiels. Cela leur donne l'opportunité, utilisée à quelques moments clés pour les représentants français, de faire front commun devant les équipes PISA-OCDE et demander par exemple des changements méthodologiques ou la mise en place d'expérimentations complémentaires pour s'assurer de la validité de certains choix faits par l'organisation. Ce fût le cas par exemple du débat lancé par le NPM français au premier cycle de l'évaluation sur la méthode de recueil de la profession des parents des élèves participant à PISA. Considérant que la méthode validée par l'OCDE n'était pas suffisamment robuste, il avait alors porté le sujet auprès de l'OCDE, avec le soutien de ses homologues japonais, allemand et tchécoslovaque en proposant une expérimentation. L'expérimentation a été acceptée par l'OCDE. Néanmoins, les résultats, bien qu'incitant selon le NPM à revoir la méthode, n'ont pas eu l'effet de convaincre l'OCDE de modifier le mode de recueil. Les NPM des autres pays participants sont d'ailleurs qualifiés de « collègues » de travail, avec lesquels les relations professionnelles se tissent et sont d'autant plus fortes que les NPM restent plusieurs années en charge de cette mission de représentation. L'exemple fourni par un autre NPM français illustrant le travail conjoint entre les représentants internationaux au moment de valider la pertinence culturelle des items témoigne bien de la perception de la constitution d'un groupe soucieux de prendre en compte les remarques de chacun :

« Comme les items sont tous ancrés dans des situations de vie réelle, le moindre contexte fait que dans un pays ou dans un autre, les connaissances des élèves peuvent relever non pas seulement des connaissances mathématiques mais d'autres dimensions culturellement très très influencées qui pourraient biaiser la mesure. On est quand même assez vigilants lorsqu'on conçoit évidemment. On ne va pas faire en sorte de favoriser tel ou tel pays et on y pense. Mais on avait par exemple proposé un item qui avait pour cadre le football. Nos collègues brésiliens nous ont

tout de suite dit que c'était impossible de faire un item sur le football parce que chez eux c'est tellement culturellement fort qu'ils avaient repéré à l'avance qu'il pouvait y avoir de très grandes différences entre les performances par exemple des filles et des garçons sur ces items-là. On a des pays qui sont parfois assez homogènes culturellement et d'autres plus différents. On fait attention à tout ça. On se parle. J'ai rencontré des gens qui travaillent sur PISA depuis des années, puisqu'il y a des NPM qui sont là depuis le début pratiquement. La collègue russe par exemple, la collègue chilienne aussi. »¹⁸⁴

Leur rôle central bien sûr pour les pays participants qui ont avec eux une oreille active dans toutes les réunions portant sur PISA et un œil sur tous les documents préparatoires produits à la fois par l'OCDE et par le consortium chargé du traitement statistique des données. Mais il est également central pour l'OCDE car les NPM sont les garants de la bonne participation des pays à l'évaluation. Si les orientations et les contenus des évaluations sont validés par les NPM, alors la légitimité des pays à remettre en question les résultats éventuellement mauvais en attaquant la méthodologie de l'étude est réduite. Afin de s'assurer que le rôle du NPM est bien compris par tous et surtout compris de la même façon, l'organisation a produit un outil reprenant à la fois leurs missions et leurs obligations sous la forme d'un manuel. Fournis par l'OCDE et réalisés par les membres du consortium chargé de l'organisation statistique des évaluations, ils sont très détaillés. La mouture 2018 signée du consortium ETS, Pearson, Westat et Capstan fait ainsi plus de 150 pages et a, selon ce qui y figure, plusieurs objectifs :

- « - Dépeindre le rôle du NPM dans le contexte international plus large de la mise en œuvre du projet ;
- Fournir des informations pratiques pour planifier et mettre en place des tâches et des informations qui ont des implications pour l'allocation des ressources nationales ;
- Fournir une vue générale du portail PISA ;
- Aider à fournir clairement aux NPM les raisons d'être de ses tâches en spécifiant le lien entre les activités des NPM et les Standards Techniques de PISA 2018 ; et

¹⁸⁴ Entretien avec Franck Salles, Chargé d'étude sur l'évaluation des élèves, National Project Manager, DEPP, 11 juillet 2017.

- Expliquer les protocoles d'une communication réussie avec les membres du consortium et les autres parties prenantes. »¹⁸⁵

Le fait que le manuel soit produit par le consortium statistique, centré sur l'organisation du travail technique, et seulement distribué par l'OCDE n'est pas neutre. En abordant le rôle du NPM sous l'angle statistique, l'OCDE le range du côté des experts techniques plus que des représentants politiques. Nous verrons par la suite que ce n'est pas exactement la lecture qu'en a eu la France qui a compté sur ses NPM pour regarder avec un œil stratégique les propositions faites par l'OCDE et le consortium, et faire remonter les éventuels points de crispation qui pouvaient émerger concernant à la fois le contenu des questions posées et la méthodologie employée. Cette méthodologie, dont le cadre général a été pensé par l'OCDE et travaillé avec le CPP, est mise en musique par le consortium statistique qui constitue le référent scientifique et technique de ce panorama d'acteurs.

iv. Un consortium statistique pour faire de PISA une évaluation incontestable

La question des outils employés est centrale pour l'OCDE qui demande, par la voix du CPP, que soit inscrit en bonne place comme objectif à PISA l'utilisation « *d'outils d'évaluation qui mettent l'accent sur l'authenticité et la validité pour l'éducation et qui autorisent des mesures fiables* », et « *de méthodes modernes à tous les stades de la collecte des données, de leur analyse et de l'établissement des rapports.* »¹⁸⁶

Ce cadrage strict est d'autant plus suivi avec attention par les pays que PISA est l'étude de loin la plus coûteuse à laquelle ils participent. Pour la France, le budget alloué à chaque cycle, hormis le coût de fonctionnement interne au ministère du fait de la constitution d'une équipe dédiée au sein de son administration, se monte à plusieurs centaines de milliers d'euros selon les

¹⁸⁵ OCDE, *PISA National project manager*, 10 mars 2017, p.13.

¹⁸⁶ DEELSA/PISA/BPC(99)9, p.5.

options choisies. En comparaison, d'autres études internationales auxquelles la France a participé ont coûté trois à dix fois moins. C'est le cas de TIMSS par exemple, qui a coûté autour de 90000 euros, ou de PIRLS, menée également par l'IEA et portant sur la compréhension de l'écrit des élèves de CM1, pour laquelle la France a dépensé autour de 40000 euros par édition. Le budget total de PISA, tous pays confondus, a été dès le départ très élevé pour une étude de ce type. Plus de 17,5 millions de francs ont été à l'époque prévus pour le premier cycle PISA. Autant dire que l'enjeu de fiabilité des données obtenues est grand et constitue un défi difficile à relever compte tenu de la complexité de ce qui est souhaité par l'OCDE. Elle le dira elle-même, les méthodes choisies sont complexes car :

« les enquêtes de grande envergure telles que celle du programme PISA se heurtent à un certain nombre de difficultés qui tiennent aux divergences d'intérêt des parties prenantes et aux difficultés concrètes de mise en œuvre de pareils travaux. S'il est par exemple souhaitable de pouvoir couvrir une large gamme de matières scolaires, il reste que le temps qu'il est possible de consacrer à l'administration des tests individuels aux élèves est dans la pratique limité. De plus, la structure des systèmes scolaires et la prise en compte des variables liées à l'établissement compliquent l'application des méthodes d'échantillonnage. »¹⁸⁷

Charge alors aux candidats statisticiens de s'inscrire au mieux dans cette exigence de robustesse technique. Un appel à candidature est lancé par l'OCDE dès octobre 1998 pour trouver l'acteur technique du grand projet PISA, l'entité responsable de construire l'évaluation sur le plan statistique. Les exigences de l'organisation, verbalisées par la voix des pays participants et matérialisées dans un cahier des charges fourni d'une centaine de pages, sont nombreuses et précises. Si une marge de manœuvre est laissée au candidat pour créer les outils de recueil et la méthodologie à employer pour constituer les échelles sur lesquelles les résultats aux tests seront positionnés, plusieurs exigences fondamentales doivent être prises en compte, et en particulier l'assurance d'une comparabilité fiable des résultats. Huit principes sont énoncés dans le cahier des charges : dans l'ordre, « fiabilité, exactitude et comparabilité », « masse critique

¹⁸⁷ DEELSA-PISA-BPC(99)9- FRE, p.19.

des indicateurs », « efficacité », « rapport coût-efficacité », « partenariat », « équilibre », « transparence ». La validité statistique des résultats apparaît comme la préoccupation première du CPP, avant même d'autres principes permettant d'assurer la qualité d'une évaluation, et notamment la transparence du processus évaluatif impliquant que chaque maillon de la chaîne soit tenu de rendre des comptes de ce qu'il produit¹⁸⁸. Nous verrons que ce principe essentiel à PISA sera au cœur de la bataille d'influence qui aura lieu entre les évaluateurs nationaux français et les équipes de PISA de l'OCDE.

Trois offres sont reçues, remises par des consortiums menés par des acteurs du monde universitaire et de la recherche, l'Université de Bourgogne en France par la voix de l'IREDU associé à la DEP, le Boston College aux Etats Unis et l'*Australian Council for Educational Research* en Australie. Elles sont passées au crible par un groupe d' « experts internationaux reconnus ¹⁸⁹ » validés par le CPP, et qui sont issus d'instituts de statistiques renommés. Si l'on regarde la composition de ce groupe, il apparaît rapidement que les pays anglo-saxons sont surreprésentés. Sur les six membres, deux sont américains, et deux sont britanniques¹⁹⁰. La prédominance de ces pays peut s'expliquer par l'importance du courant de psychométrie dans leur sphère universitaire et par l'expérience qu'ils ont au niveau national de l'utilisation d'évaluations standardisées pour mesurer les résultats de leurs élèves. Cela vaudra toutefois à PISA, dès le départ, d'être qualifiée par les évaluateurs français d' « évaluation à l'américaine »,

¹⁸⁸ La transparence est, pour les évaluateurs français, un principe clé assurant la validité d'une évaluation. Elle est notamment énoncée dans la Charte de l'évaluation des politiques publiques et des programmes publics adoptée en juin 2006 par les membres de la Société Française de l'Evaluation aux côtés de six autres principes que sont la pluralité, la distanciation, la compétence, le respect des personnes, l'opportunité et la responsabilité. La SFE a été fondée à la fin des années 1990 et a pour objectif selon ses statuts de « *contribuer au développement de l'évaluation et de promouvoir son utilisation dans les organisations publiques et privées. (...) Animée par un réseau de bénévoles, elle rassemble des professionnels, chercheurs, agents de la fonction publique et consultants. (...) La SFE a défini les principes spécifiques de l'Evaluation dans une Charte adoptée en 2006 qui constitue une référence pour les acteurs intervenant dans ce domaine* ».

¹⁸⁹ DEELSA-PISA-BPC(98)3-ENG, p. 4.

¹⁹⁰ Les experts sont Nadir Atash (PARSA, Etats-Unis), Jürgen Baumert (Max Planck Institute for Human Development and Education, Allemagne), Marilyn Binkley (National Center for Education Statistics, Etats-Unis), David Hawker (Qualification and Curriculum Authority, Royaume-Uni), Neville Postlethwaite (International Consultant, Royaume-Uni/France) and Jean-Paul Reeff (Ministry of Education, Luxembourg) .

signifiant clairement qu'elle ne serait a priori pas adaptée au système français et donc peu représentative des résultats réels des élèves¹⁹¹. Nous verrons plus tard que toute la question réside dans ce que l'on entend par « résultat », reformulé par l'OCDE par le concept de « performance », et dont les définitions divergent.

Les experts se prononcent finalement sur le choix du consortium mené par ACER, selon évidemment des critères de compétences techniques de la proposition faite, mais aussi selon des critères d'indépendances des membres du consortium par rapport aux ministères des pays participants. La DEPP faisait partie du consortium mené par l'IREDU et cela a joué en sa défaveur.

Les méthodes de communication anglo-saxonnes ont également joué un rôle non négligeable dans le choix final en « montrant » le sérieux des membres du consortium ACER aux experts et aux membres du CPP. Les français de la DEPP présents alors se souviennent de l'image de maîtrise qui se dégageait de la présentation faite par ACER de sa proposition, en comparaison des autres.

« Il y avait eu une réponse française à l'appel d'offre, faite par la DEPP. C'est vrai que quand on voit le document qui a été produit, c'était assez amateur. J'avais été surpris par le professionnalisme d'ACER. Maintenant tout le monde fait ça, mais il y a 20 ans c'était nouveau, des belles photos, un argumentaire fluide, ils disaient qu'ils allaient faire des trucs super. C'était de la super com'. Les anglo-saxons ont la capacité à communiquer. Sur le plan technique, le document n'était pas plus technique que les autres. »¹⁹²

La DEPP avait par ailleurs à l'époque peu l'expérience de répondre à des appels d'offre de ce type et son manque d'habitude, au-delà des défauts « marketing » de sa proposition, a pu lui porter préjudice.

ACER est donc chargé de mener l'évaluation sur le plan statistique. Il sera remplacé en 2015 par *Education testing service* (ETS), un autre anglo-saxon déjà membre du consortium précédant et connu pour avoir en charge les évaluations TOELF dans plus de 180 pays. Ceux qui seront dénommés plus simplement « le

¹⁹¹ Voir chapitre 4 pour plus de détails.

¹⁹² Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 21 mars 2019.

consortium » deviendront les interlocuteurs à la fois de l'OCDE pour préparer chaque cycle d'évaluation mais aussi des statisticiens nationaux chargés de suivre les évaluations PISA. L'OCDE se pose en contrôleur de l'activité du consortium.

Les règles qui régissent leurs relations sont ainsi décrites :

« Dans le cadre de ses activités de suivi, le Secrétariat s'efforce de savoir dans quelle mesure le consortium a réussi à :

- respecter au mieux les délais impartis ;
- établir des comptes rendus détaillés et une documentation concernant les réunions et les communications ;
- promouvoir le consensus et le maximum de concorde entre les pays ;
- remettre au Secrétariat de l'OCDE des rapports complets sur les principes de gestion prévus ;
- coordonner les activités de projet menées par les différents membres du consortium et les groupes d'experts fonctionnels ;
- établir des méthodes de communication efficaces entre tous les participants ;
- dresser des plans pour les imprévus. »¹⁹³

Le consortium intervient tout au long de la période de trois ans qui caractérise chaque cycle. Il est notamment chargé du tirage des échantillons d'établissements participants à l'évaluation, l'échantillonnage des élèves étant laissé à la charge des statisticiens nationaux afin de garantir le plein anonymat des résultats traités ensuite. Cet échantillonnage est fait deux fois, pour la phase de test organisée systématiquement et pour la véritable évaluation. Le consortium récupère ensuite les bases de données nettoyées par les statisticiens nationaux, ceux de la DEPP dans le cas de la France. Ils se chargent ensuite de réaliser tous les traitements de données pour aboutir à des valeurs correspondant aux résultats des élèves et des établissements, pouvant être comparées entre elles et positionnées sur des échelles permettant de les classer. Les données finales sont ensuite transmises à l'OCDE qui en fait l'analyse sous la forme d'indicateurs qu'elle crée et qu'elle publie ensuite. Les statisticiens nationaux n'ont finalement qu'un rôle réduit entre le moment où ils transmettent les données brutes nettoyées et le moment où l'OCDE les a en main, si ce n'est de répondre aux questions du consortium sur

¹⁹³ DEELSA/PISA/BPC(99)9, p.20.

d'éventuelles données manquantes ou présentant des anomalies.

On le voit, l'OCDE a pensé l'évaluation PISA en faisant intervenir un nombre d'acteurs important à la fois politiques et techniques, à la fois nationaux et internationaux. La règle du consensus est posée dès le départ et est en théorie appliquée à tous les niveaux de prise de décision. Ce consensus est d'autant plus nécessaire pour l'organisation que le programme PISA est construit selon des principes totalement nouveaux dans le domaine de l'évaluation des politiques éducatives et bouscule les habitudes des pays en la matière. Ces derniers, pour lesquels les évaluations nationales font déjà souvent l'objet d'oppositions politiques et de débats techniques, doivent s'engager sur un temps long dans une évaluation qui sort de la classique mesure des acquis d'un programme scolaire pour s'engager sur la piste – aux résultats impossibles à anticiper – de l'acquisition de compétences transversales au cours de la scolarité. PISA devient pour eux un nouvel instrument d'action publique.

3) La nouveauté PISA : un instrument d'action publique inédit pour évaluer les résultats de l'éducation

L'évaluation PISA bénéficie d'une aura particulière dans le monde des évaluations internationales, y compris au sein même des évaluations menées par l'OCDE. Bien d'autres projets ont été lancés ces dernières décennies mais n'ont pas reçu le même accueil médiatique et politique, bien qu'ils traitent de problématiques éducatives ayant un impact tout aussi direct sur l'état de l'économie et pouvant se rapporter au sujet cher à l'OCDE du « capital humain ». On peut citer parmi elles le PIAAC, le Programme pour l'évaluation internationale des compétences des adultes¹⁹⁴, lancé en 2011, qui mesure cette fois la maîtrise par les adultes de compétences clés (lecture, mathématique, résolution de problème) et rassemble des données sur la façon dont les

¹⁹⁴ En anglais Programme for the International Assessment of Adult Competencies.

compétences sont utilisées à la maison, au travail, et plus largement dans la vie quotidienne. Mené dans plus de 40 pays et économies, il a pour objectif, selon l'OCDE, de permettre aux participants de recueillir des données permettant de travailler à l'acquisition de compétences nécessaires « pour que les individus participent à la vie de la société et pour que les économies prospèrent »¹⁹⁵. L'objectif final de l'étude est donc fortement similaire à celui de PISA, si ce n'est qu'il s'intéresse aux adultes et non aux élèves en cours de scolarité obligatoire. Pourtant, PIAAC n'a que peu d'écho dans les médias et n'est pas vraiment repris publiquement par les décideurs publics. En France, l'évaluation est du ressort du ministère du travail et non du ministère de l'éducation, ce qui explique peut-être les moindres passions que ce programme suscite.

D'autres évaluations de l'OCDE, portant pourtant sur l'enseignement obligatoire, ne font pas davantage l'objet, en France, d'une intense diffusion publique lorsqu'il s'agit de construire les politiques éducatives à venir. L'évaluation TALIS notamment, qui traite du sujet central pour la politique éducative de l'environnement d'apprentissage des élèves sous l'angle des conditions de travail des enseignants et des chefs d'établissement, reste réservée au cercle interne de l'administration et des cabinets ministériels¹⁹⁶. Et cela alors qu'elle interroge plus de 260 000 enseignants dans 15000 écoles de plus de 48 pays et économies.

Le crédit d'« expert » souvent alloué à l'OCDE comme organisation internationale tentaculaire n'explique donc pas à lui seul pourquoi PISA reste l'évaluation internationale la plus reprise, à la fois médiatiquement et au sein des administrations travaillant sur les questions d'éducation, en France. Nous faisons l'hypothèse qu'il faut chercher non seulement du côté du fond de l'évaluation mais aussi de la forme que revêt la publication des résultats.

¹⁹⁵ <http://www.oecd.org/skills/piaac/>

¹⁹⁶ Entretien avec Agathe Cagé, Conseillère au cabinet du ministre de l'Education Nationale en charge du second degré, des programmes et de l'évaluation des élèves, octobre 2016.

a) Une évaluation novatrice ? L'inspiration de TIMMS

Si PISA est reconnue comme ayant défini de nouveaux outils statistiques pour évaluer les politiques éducatives, elle tire toutefois son inspiration d'études ambitieuses déjà menées par d'autres organisations internationales, que ses concepteurs visent à dépasser. Les évaluations de l'Association Internationale pour l'Evaluation du Rendement Scolaire (IEA) ont été en particulier une source d'idées pour PISA, venues avec Andreas Schleicher lors de son entrée à l'OCDE.

L'IAE est au début des années 1990 la référence des organisations internationales indépendantes travaillant à l'évaluation des résultats de l'éducation. Elle s'est constituée, tout comme nombre d'organisations internationales travaillant dans la période d'après-deuxième guerre mondiale à la fin des années 1950. Elle est créée dans le sillage des activités menées par l'Unesco en matière de politique éducative, par le biais de son institut dédié, l'Institut d'Education. Le groupe de pays fondateurs souhaitait alors identifier les facteurs favorisant un bon « rendement éducatif »¹⁹⁷ en mettant en place une évaluation commune permettant de donner aux résultats obtenus un point-de-vue comparatif. Impulsée par des chercheurs d'une douzaine de pays différents, cette organisation internationale s'engage dans une méthodologie nouvelle: celle des *surveys* normatifs. L'objectif est alors d'évaluer par une méthode scientifique la recevabilité des hypothèses concernant les forces et les faiblesses des systèmes éducatifs¹⁹⁸.

L'OCDE fait explicitement référence à ces évaluations, et à une en particulier nommée TIMSS (Trends in International Mathematics and Science Study), comme à un exemple à « éventuellement » suivre dans le cadre stratégique qu'elle publie au début des travaux de construction de l'évaluation :

« Pour le premier cycle, la lecture dans la langue d'enseignement sera la

¹⁹⁷ Alejandro Tiana Ferrer, « Pilotage par les résultats et amélioration de l'éducation : deux réalités indissociables ? », *Administration et éducation*, n°98, 2003. Il fût président de l'IEA de 1999 à 2004.

¹⁹⁸ Marcel Crahay et Arlette Delhaxe, *art. cit.* 2004.

matière dominante, et les mathématiques et les sciences les matières secondaires ; on s'inspirera peut-être à ce propos des travaux conduits dans le cadre de la Troisième étude internationale de mathématiques et des sciences (TIMSS) menée par l'Association Internationale pour l'Évaluation des Résultats Scolaires (IEA). » ¹⁹⁹

Les travaux de l'IEA et de l'OCDE dans le champ éducatif sont souvent allés de pair, l'IEA étant le précurseur sur certains types d'indicateurs, et en particulier ceux allant au-delà des données contextuelles et systémiques de l'éducation. Ainsi, la publication de l'OCDE *Regards sur l'éducation* utilisait dans ses premières éditions les indicateurs de l'IEA sur les résultats éducatifs. De la même façon, l'évaluation TIMSS a inspiré la construction de PISA, en étant adaptée « à la sauce » OCDE par l'intermédiaire du transmetteur Andreas Schleicher.

L'évaluation TIMSS a été lancée pour la première fois en 1995. Si l'éducation comparée, permettant par des données chiffrées d'étudier quantitativement plusieurs systèmes éducatifs en parallèle, était déjà un sujet d'intérêt pour la recherche et l'évaluation des politiques éducatives depuis le début du XIXe siècle²⁰⁰ (et plus encore au cours du XXe siècle via notamment toutes les études menées par les différentes instances internationales mentionnée précédemment) c'est avec les travaux de l'IAE et TIMSS en particulier que les indicateurs internationaux du rendement scolaire tels que nous les connaissons aujourd'hui voient le jour. Il s'agit alors davantage de trouver des indices pour répondre aux problématiques communes des systèmes éducatifs de pays partenaires économiques que de réfléchir au fonctionnement de différents systèmes éducatifs.

TIMSS est une évaluation cyclique menée tous les quatre ans qui étudie la performance des élèves de plusieurs niveaux éducatifs en sciences et mathématiques. Elle a porté la première année sur 41 pays. L'évaluation PISA

¹⁹⁹ DEELSA/ED/CERI/CD(97)6-FRE, p. 9.

²⁰⁰ Marc-Antoine Jullien, articles du « journal d'éducation », 1817, donne la définition de l'éducation comparée, dans Dominique Groux, « L'éducation comparée : approches actuelles et perspectives de développement », *Revue française de pédagogie*, volume 121, 1997.

présente en essence de nombreux points communs avec TIMSS. Elle est cyclique elle-aussi, cette fois tous les trois ans, et permet donc une comparaison dans le temps des résultats des pays. Les pays participants sont au départ similaires, les pays européens et anglo-saxons, mais aussi certains pays d'Asie comme Singapour, la Corée du Sud ou le Japon. Toutefois l'OCDE a voulu se démarquer de ce qui avait été mis en place pour TIMSS sur plusieurs aspects fondamentaux de son évaluation et aller plus loin dans l'établissement de corrélations qu'elle veut explicatives sur les résultats des élèves.

Les deux évaluations diffèrent également sur le choix de ce qui est évalué : alors que TIMSS basait ses items sur le plus petit dénominateur commun des programmes d'enseignement dans tous les pays participants, PISA au contraire les construit en prenant l'ensemble des programmes, même si tous les élèves n'ont pas reçu la totalité des enseignements.

Cette volonté de se détacher du programme a nécessité un travail fin de conceptualisation de l'objet de l'évaluation. Si l'on n'évalue pas les acquis à partir d'une liste de connaissances prédéterminée et qu'il convient d'apprendre, comment évaluer les résultats de l'éducation ? PISA répond en s'intéressant à la performance des élèves dans des situations de la vie quotidienne nécessitant de mobiliser des connaissances apprises en classe, peu importe le niveau. Et cet éloignement de l'objet classique des évaluations scolaires a impliqué de réfléchir à de nouveaux outils de mesure et de rendu des analyses des résultats. Sur le fond d'abord, la méthodologie est novatrice tant du point-de-vue de la population cible que de ce qui est mesuré. Sur la forme, la périodicité du rendu des résultats a contribué à faire de PISA un programme d'évaluation incontournable au sens premier du terme, qui ne peut pas être évité.

b) L'introduction de concepts inédits et de méthodologies nouvelles

i. Les théories du capital humain à l'origine du changement de paradigme de l'OCDE

Alors que l'OCDE est reconnue internationalement pour son expertise économique qui prône des réformes néolibérales, elle semble faire un pas de côté dans le domaine de l'éducation en travaillant sur le capital humain et son apport dans la réduction des écarts de développement au niveau mondial et dans la constitution d'une « société de la connaissance »²⁰¹. Les publications du CERI énoncent la nécessité de travailler dans cette direction : « *Dans un monde de plus en plus intégré, qui se transforme rapidement, la réussite des nations, des collectivités et des individus peut être plus que jamais liée à leur capacité de changer, d'apprendre et de partager leurs connaissances.* »²⁰²

En travaillant sur le capital humain, elle fait sienne et applique à l'évaluation des politiques éducatives un concept issu de l'économie du développement apparu dans les années 1960 et théorisé par le prix Nobel d'économie Gary Becker au début des années 1990. Les individus, loin de se résumer à de simples facteurs de productions, acquièrent des compétences tout au long de leur formation et les valorisent sur le marché du travail. Au-delà du niveau individuel, le capital humain est alors un facteur de croissance économique²⁰³. C'est bien dans cette perspective que l'OCDE travaille sur le sujet, en liant croissance économique et développement du capital humain par l'éducation. PISA s'inscrit complètement dans cette vision d'un « monde qui change », et qui implique d'envisager l'éducation comme un processus permettant d'acquérir non pas un stock de connaissances mais bien un flux de compétences sans cesse enrichi favorisant l'adaptabilité des individus à

²⁰¹ Annick Kieffer et Rémi Tréhin-Lalanne, *art cit.*

²⁰² DEELSA/ELSA/ED/CERI/CD(2000)3/REV1.

²⁰³ Gary Becker, *Human Capital : A Theoretical and Empirical Analysis, with Special Reference to Education*, University of Chicago Press ; voir aussi Mankiw G., Romer D. et Weil D., « A Contribution to the Empirics of Economic Growth », *Quarterly Journal of Economics*, 1992.

l'évolution du monde, et du monde du travail.

Cette compréhension de l'objectif de l'éducation induit, pour PISA, la mise en place d'une méthodologie spécifique jamais vue jusqu'alors. La population évaluée et le contenu de l'évaluation sont les deux paramètres qui font de PISA une évaluation novatrice.

ii. Un échantillon constitué d'élèves d'une même classe d'âge

Si la méthode de recueil a été proposée par le consortium retenu pour l'évaluation, plusieurs grands principes faisant toute la nouveauté de PISA étaient bien présents dans le cahier des charges produit par l'OCDE. L'un d'entre eux était tout à fait inédit et très différent de toutes les autres évaluations jusqu'alors, qu'elles aient été nationales ou internationales. Il s'agit du principe de s'adresser aux élèves d'une même classe d'âge et non pas aux élèves d'un même niveau scolaire. L'âge en question est même précisé dans le cahier des charges, « entre 15 et 16 ans », pour correspondre à l'âge moyen de la fin de la scolarité obligatoire dans la plupart des pays participants. Pour la France, cela signifie que les élèves sont choisis majoritairement à la fois dans les classes de 3^e, dernière année de collège, et de Seconde, première année de lycée, avec donc des écarts importants en termes d'acquisition des connaissances du programme. Des élèves ayant redoublé et d'autres ayant sauté une classe peuvent également être intégrés dans l'échantillon.

Cette nouveauté en matière d'échantillonnage, véritable grand écart méthodologique, a pris presque 5 ans à être accepté par tous les pays participants à PISA et à être intégré dans le cahier des charges de l'évaluation. C'est bien là la force méthodologique de PISA, désormais collectivement reconnue, y compris par les évaluateurs français qui étaient pour certains peu enthousiastes à participer à des études qui n'évaluaient pas le même contenu d'un pays à un autre tout en se voulant comparatiste. C'était notamment le cas au sein de la DEPP en France. Des mots d'un ancien NPM présent à l'époque du lancement de PISA :

« PISA c'était une façon de concurrencer TIMSS, avec une autre approche. L'IAE travaillait en testant une classe, un niveau, au lycée et en primaire. Elle se fonde sur les curriculums, et on ne va pas interroger les gens qui ne sont pas dans le curriculum. Les élèves ont des âges différents parce qu'on est dans une approche par niveau. PISA a été destiné à surmonter tout ça. C'était enthousiasmant. Avec TIMSS il y avait des problèmes, et on s'est dit que PISA allait les résoudre. »²⁰⁴

Pour ceux qui souhaitaient tout de même rester plus proche des évaluations connues, une option payante était offerte, celle d'obtenir aussi des évaluations basées sur l'année d'études. En pratique, elle n'a pas réellement été demandée par les participants à PISA. La France devrait toutefois recevoir des données par niveau scolaire lors du prochain cycle prévu en 2022²⁰⁵.

Si l'on décide de ne plus évaluer les élèves d'un même niveau scolaire, alors il est impensable de fonder les items des tests sur l'acquisition du contenu d'un programme. L'OCDE décide donc de travailler sur la littératie, en accord avec sa conception de l'éducation tout au long de la vie, permettant d'acquérir non pas des connaissances sèches, sorties d'un contexte, mais bien des compétences utilisables dans la vie de tous les jours pour résoudre toutes les tâches auxquels les individus peuvent être confrontés.

iii. Une nouvelle définition des résultats de l'éducation par la performance

Ces jeunes de 15 ans sont testés sur leur « *literacy* » un concept difficile à traduire en français, parfois traduit par « culture », parfois traduit par l'anglicisme littératie. L'OCDE en donne une définition dans les derniers volumes de PISA 2018, après plusieurs détours : la littératie correspond à « *la capacité des élèves d'exploiter des savoirs et savoir-faire dans des domaines clés et d'analyser, de raisonner et*

²⁰⁴ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 18 mars 2019.

²⁰⁵ Du fait de la crise sanitaire du covid-19, le cycle PISA 2021 a été reporté d'une année, de nombreux pays n'ayant pas pu réaliser la phase de test prévue en mai 2020 à cause du confinement.

*de communiquer lorsqu'ils énoncent, résolvent et interprètent des problèmes qui s'inscrivent dans divers contextes. »*²⁰⁶

Dans le premier volume de PISA qui reprend les fondements conceptuels de l'évaluation, l'OCDE énonce la raison d'être de ce choix d'évaluer la littératie et non l'acquisition de connaissances :

« Le concept PISA de littératie ou de culture (compréhension de l'écrit, culture mathématique et culture scientifique) est beaucoup plus vaste que la notion historique d'aptitude à lire et écrire. A présent, il est communément accepté qu'une personne instruite dispose de toute une gamme de compétences. Par ailleurs, on ne considère plus qu'il existe un seuil spécifique séparant les individus instruits et les illettrés. La « culture », l'« instruction » se mesurent sur un continuum, une échelle de niveaux de compétence successifs, et non pas comme quelque chose dont on dispose ou on ne dispose pas. Dans PISA, il s'agit des connaissances et savoir-faire dont on a besoin dans la vie d'adulte. Leur acquisition est un processus d'apprentissage qui se poursuit tout au long de la vie, et pas seulement à l'école ou dans l'enseignement formel : on apprend aussi à travers des interactions avec des jeunes du même âge, des collègues de travail et la communauté élargie. »²⁰⁷

Grande nouveauté, PISA s'intéresse aux acquis des élèves et est considéré par tous comme une évaluation des résultats de l'éducation, mais de l'éducation au sens large, pas seulement fournie par les systèmes éducatifs. Toujours est-il que l'évaluation portera bien sur les résultats des élèves de 15 ans après plusieurs années dans un système scolaire et sera comprise comme telle par les décideurs publics et par tous ceux qui en utiliseront les résultats.

Avec ce nouvel outil d'action publique il ne s'agit donc pas seulement pour les décideurs de regarder le niveau d'acquisition des élèves de leur pays, mais bien la capacité de mobilisation des connaissances, de « *mise en œuvre des savoirs et savoir-faire dans des tâches et des défis quotidiens, que ce soit en famille ou dans le monde du travail.* »²⁰⁸ La distinction entre « tâches » et « défis » est intéressante car elle

²⁰⁶ OCDE, *Résultats de PISA 2018 (Volume I) : Savoirs et savoir-faire des élèves*, Editions OCDE, 2019, p.26.

²⁰⁷ OCDE, *PISA 2000, Volume 1*, Editions OCDE, 2001, p.5.

²⁰⁸ Thierry Rocher, *PISA, une belle enquête : lire attentivement la notice*. Association Française des Acteurs de l'Éducation, « Administration & Education, 1, n°145, 2015.

présuppose que PISA va bien plus loin que ce qui peut être attendu en classe en termes de maîtrise, mais se donne pour objectif de plonger les élèves, au moment des tests, dans des situations inconnues et nécessitant de construire des réponses nouvelles plutôt que de mobiliser par la répétition des informations reçues dans leur cursus scolaire.

C'est là que toute la complexité de l'objet de ces évaluations apparaît. Les connaissances et les qualifications ont déjà été travaillées au niveau international et national, notamment au niveau national par les pays via la définition de leurs programmes scolaires d'enseignement et par la mise en place de certifications sanctionnant l'atteinte de différents niveaux scolaires paliers. Nous pourrions définir ici les connaissances comme l'ensemble du contenu scolaire appris, retenu et utilisé dans le cadre scolaire par un individu au cours de sa formation, à partir donc d'un programme scolaire pré-déterminé. Dans le cas français, ce programme est défini au niveau national et approuvé par le ministre de l'éducation nationale²⁰⁹. Le concept de connaissances s'inscrit dans le champ scolaire, à la différence des qualifications qui se rapportent davantage aux capacités de travail des individus valables dans le champ de l'emploi. Ce sont ces connaissances qui ont fait par le passé – et de nouveau avec le ministère de Jean-Michel Blanquer - l'objet d'évaluations exhaustives avec notamment celles mises en place par Lionel Jospin.

Les compétences quant à elle sont, dans les années 1990, un sujet de travail très nouveau, pour l'OCDE comme pour les pays membres. Si les théories du capital humain sont une inspiration forte pour la construction du cadre théorique de PISA, l'organisation n'est pas encore pleinement engagée sur ce sujet. Le cadre général est toujours majoritairement économique, même si la définition des compétences donnée par l'OCDE donne une place centrale à la formation au sens large, aussi bien initiale que continue. Les compétences sont définies

²⁰⁹ Conçu jusqu'en 1989 par l'Inspection générale, il est remis entre les mains du Conseil national des programmes par la loi d'orientation de 1989 sur recommandations du rapport du Collège de France « Pour un enseignement de l'avenir ». A partir de cette date, des groupes d'experts disciplinaires sont chargés de proposer un projet de programme ensuite soumis au CNP.

précisément par l'OCDE dans un document de 2011 présenté aux ministres de l'éducation. Elles sont entendues comme « *l'ensemble des connaissances, caractéristiques et aptitudes qui permettent à un individu de mener systématiquement à bien une tâche ou une activité, au sens large ou strict, et qui peuvent être renforcées et développées par le biais de la formation.* »²¹⁰ Elles sont toutefois invoquées dans le contexte de la concurrence commerciale accrue entre les pays et les taux de chômage élevés des pays de l'OCDE en particulier après la crise de 2008. L'objectif d'usage de la compétence, la réalisation d'une tâche ou d'une activité grâce à la formation, est bien celui évalué par PISA, appliqué aux élèves en fin de scolarité obligatoire.

Si les publications et les documents de travail parlent parfois de résultats des élèves ou de rendement de l'éducation pour parler de l'évaluation de ces compétences, un concept clé est introduit dans le langage des politiques éducatives par PISA : celui de « performance ». Absent du vocabulaire français des politiques éducatives jusqu'alors, il est employé par l'OCDE pour parler des résultats PISA dès la conception de l'évaluation. Issu de l'anglais « to perform » qui signifie « réaliser, accomplir », emprunté au lexique du sport et ayant infusé dans le champ des politiques publiques en commençant par les politiques économiques, le concept de performance est déjà utilisé pour parler des résultats de l'éducation dans les pays anglo-saxons depuis des décennies lorsque PISA est conçu. Les britanniques avaient notamment déjà mis en place dans les années 1970 une unité d'évaluation de la performance, the Assessment of Performance Unit, au sein de leur Département de l'éducation et des sciences afin de promouvoir l'évaluation des résultats et de la performance des élèves. Impliquant un management par les résultats et plus par les moyens, la performance est dès le départ posée en objet évaluatif par les concepteurs de PISA, dès les premières réunions. Lors des échanges sur la conception des échelles d'évaluation, ils parlent bien non pas de niveaux de résultats mais de niveaux de performance. « (...) *La détermination de niveaux de performance définis empiriquement requière le*

²¹⁰ OCDE, *Elaborer une stratégie en faveur des compétences*, réunion du Conseil au niveau des ministres, Paris, 25 et 26 mai 2011, p.5.

développement d'items suffisants dans chaque catégorie », énonce par exemple l'OCDE au cours des discussions sur la création des questionnaires donnés aux élèves participant à PISA.

Ce concept de performance, nouveau dans le champ éducatif, entre peu à peu dans le langage des politiques éducatives, y compris françaises. Il faut attendre les années 2000²¹¹ pour que le concept y soit importé. Est-ce l'inspiration directe de PISA ? Ou PISA profite-t-il du courant favorable au Nouveau Management Public en France pour transposer ses concepts clés dans les politiques d'évaluation nationales, concepts qui auraient de toute façon été utilisés tôt ou tard ? Difficile à dire tant tout cela est concomitant. Toutefois, les experts français chargés alors des évaluations des résultats des élèves au niveau national que nous avons rencontrés, de la DEPP en particulier, reconnaissent le rôle qu'a eu PISA dans la banalisation de ce concept de performance utilisé pour parler des résultats de la politique éducative. Le sens accordé au mot n'est toutefois pas tout à fait le même pour les décideurs publics français et pour les évaluateurs de l'OCDE. Qu'est-ce que la performance d'un élève ? Qu'est-ce que la performance d'un système éducatif ? Pour les experts français, la performance correspond bien souvent à la conception classique de l'objet des évaluations françaises : des élèves performants sont ceux qui ont les meilleurs résultats, c'est-à-dire qui ont acquis les connaissances attendues prévues dans le programme scolaire qu'ils ont suivi. Un système éducatif performant serait par ailleurs, en toute logique, celui dans lequel les élèves seraient les plus performants. Pour les concepteurs de PISA, la performance est perçue différemment on l'a vu. Plutôt que l'acquisition d'un contenu de connaissance, il est attendu des élèves la maîtrise de compétences clés pour leur vie d'adulte. Cette différence d'acception du mot sera, on le verra, à l'origine de nombreux écarts d'interprétation des résultats PISA entre les évaluateurs français et les équipes PISA-OCDE, et à l'origine de critiques méthodologiques fortes de la part des premiers.

²¹¹ Yves Dutercq, Sylvie Cuculou, Christian Maroy, *L'école à l'épreuve de la performance*, De Boeck Supérieur, 2013.

c) Une évaluation cyclique mais distillée en continu

Du point-de-vue du fond on l'a vu, PISA est une évaluation inédite pour les concepts qu'elle utilise et la méthodologie qu'elle emploie. Du point-de-vue de la forme ensuite, l'OCDE utilise le format du rapport à ramifications multiples pour se rappeler sans cesse au souvenir des pays et faire de PISA l'incontournable en continu, bien que les évaluations en elles-mêmes aient lieu tous les trois ans. Cela a été prévu dès le départ et formulé dans le cahier des charges de l'étude comme un prérequis à l'utilisation par les décideurs publics des résultats obtenus pour l'élaboration de leur politique éducative nationale. L'OCDE est la détentrice des résultats qui sont d'abord publiés par elle, et mis à disposition des pays une fois sortis.

« En recueillant les données tous les trois ans, on devrait obtenir un apport ininterrompu de renseignements sur les acquis des élèves qui permettra d'alimenter la réflexion des pouvoirs publics, et suffisamment d'informations à verser dans les éditions annuelles des publications de l'OCDE, lesquelles continueront d'en être les principaux vecteurs de présentation. »²¹²

La forme choisie pour rendre compte des résultats de PISA est représentative d'une stratégie de communication politique dans la droite ligne des principes du « soft power ». L'OCDE n'a pas de pouvoir coercitif ou normatif, mais elle se rend incontournable par la place qu'elle occupe en termes de fréquence de publication de recommandations en matière d'orientation des politiques éducatives aussi bien d'un point-de-vue global que du point-de-vue des pays pris individuellement. Alors que les tests ont lieu tous les trois ans, ils donnent naissance à une multitude de rapports étalés dans le temps. L'évaluation cyclique et son thème principal donne lieu à un rapport puis à une série de rapports thématiques tout au long des trois années de « creux » entre deux évaluations. Grâce à eux, l'OCDE ne se fait pas oublier et le suspense de la sortie de résultats inédits se répète. Les indicateurs de PISA sont par ailleurs repris dans

²¹² DEELSA-ED-CERI-CD(97)6-FRE, p.9.

la publication annuelle *Regards sur l'éducation* et constituent, à la fin de la décennie 2010, une partie entière de ce rapport phare très fourni. Les sujets abordés par les rapports annuels sont déterminés en accord avec les pays, souvent sous l'impulsion de l'équipe PISA OCDE. Une note est par ailleurs produite tous les mois. Appelée « PISA à la loupe », elle extrait des résultats PISA pour donner des éléments de réflexion sur une thématique précise « dans l'ère du temps ». Très synthétique – moins d'une dizaine de pages -, elle s'adresse directement aux décideurs publics et dépasse le simple rendu neutre des résultats aux évaluations des volumes principaux. Des sujets faisant parfois polémiques au niveau national sont abordés. Pêle-mêle par exemple ces dernières années, le lien entre la mixité sociale à l'école et l'équité des résultats d'apprentissage, la question du choix par les filles des carrières scientifiques, le lien entre les politiques de libre choix des établissements et la mixité sociale à l'école, ou encore la prégnance de l'écologie chez les jeunes de 15 ans²¹³.

Par ailleurs, si l'évaluation est co-produite par toute la série d'acteurs nationaux et les consortiums de techniciens coordonnés par l'OCDE, le rapport final est publié sous le label OCDE. Ce mode opératoire n'est pas l'apanage de la direction de l'éducation mais se retrouve dans diverses publications de l'OCDE, jusqu'à celles faisant référence dans son domaine de prédilection : la coopération économique. Le rapport *Etude sur l'emploi* par exemple, dont le processus de rédaction et l'objet « rapport » dans l'instauration du rôle d'expert international de l'OCDE en matière de réflexion sur les politiques publiques de lutte contre le chômage ont été analysés par Vincent Gayon, emprunte le chemin de la co-écriture pour aboutir à un rapport qui est, là aussi, celui de l'OCDE²¹⁴.

²¹³ Pauline Givord, « Le renforcement de la mixité sociale à l'école a-t-il un impact sur l'équité des résultats d'apprentissage ? », *PISA à la loupe* N° 97, 11 juin 2019, et « Quels liens entre les politiques de libre choix des établissements et la mixité sociale à l'école ? », *PISA à la loupe* N°96, 15 mai 2019 ; Tarek Mostafa, « Pourquoi les filles ne sont-elles pas plus nombreuses à choisir une carrière scientifique ? », *PISA à la loupe* N°93, 17 septembre 2019 ; Alfonso Echazarra, « Les jeunes de 15 ans sont-ils devenus plus « verts » ? », *PISA à la loupe* N°87, 28 novembre 2018.

²¹⁴ Vincent Gayon, « Un atelier d'écriture internationale : l'OCDE au travail. Éléments de sociologie de la forme « rapport » », *Sociologie du travail*, Vol. 51 - n° 3, 2009.

La co-écriture du rapport final et des rapports intermédiaires est finalement l'étape à laquelle les pays sont les moins impliqués. Elle reste réalisée en premier lieu par les équipes de PISA-OCDE, puis fait l'objet d'une lecture-validation par les équipes nationales avant la publication du rapport. Chaque pays garde le droit de regard sur ses données nationales. PISA jouit toutefois d'un traitement spécial par rapport à d'autres publications phares de EDU. S'il n'est ainsi pas rare que certains demandent à l'organisation de retirer leurs données de rapports comme *Regards sur l'éducation*, généralement lorsqu'ils considèrent que la méthodologie finalement employée ne permet pas de représenter leur système - la France est particulièrement habituée à retirer ses données, exerçant par ce biais un moyen de pression sur l'organisation – c'est peu le cas sur PISA. La force statistique de l'évaluation et la méthode de co-construction des cycles voulue et orchestrée par l'OCDE est un garde-fou solide aux réclamations a posteriori des pays qui ne se sentent plus légitimes de revenir sur des choix qu'ils ont fait par le passé une fois les résultats produits, aussi mauvais soient-ils.

Conclusion du chapitre

L'OCDE fait donc irruption à la fin des années 1990 dans le champ des politiques éducatives en se dotant d'un outil puissant, PISA. Elle devient un acteur incontournable dans la lutte pour l'appropriation d'un type de pouvoir, le pouvoir scolaire, entendu comme l'espace des positions sociales à partir desquelles peut s'exercer un pouvoir sur le système d'enseignement²¹⁵.

Pour cela, elle s'appuie notamment sur l'adoption par les pays, et par la France, des préceptes du Nouveau Management Public qu'elle a par ailleurs très largement contribué à disséminer au sein de ses pays membres. Sa crédibilité n'a

²¹⁵ Par transposition à partir de la définition donnée par Pierre Clément du champ du pouvoir : « Celui-ci est en effet défini comme le lieu où s'affrontent les détenteurs des différentes espèces de pouvoirs et de capitaux, c'est-à-dire comme « *champ des rapports de force entre les pouvoirs capables de s'exercer dans les différents champs et comme champ de lutte pour transformer ces rapports de force* » (Bourdieu 2011 : 128) », dans *Réformer les programmes pour changer l'école. Une sociologie historique du champ du pouvoir scolaire*, thèse pour le doctorat de sociologie sous la direction de Bertrand Geay et soutenue à l'Université de Picardie Jules Verne le 11 mars 2013.

semble-t-il pas été réellement remise en cause par la fragilisation empirique des recommandations libérales formulées dans les années 1990 par l'OCDE au moment de l'augmentation forte du chômage dans ses pays membres dans la seconde moitié de la décennie. Elle l'a cependant forcé à se repositionner et à investir davantage la « question sociale »²¹⁶, en introduisant dans les réflexions économiques de nouvelles dimensions incitant à s'intéresser au champ éducatif, pourvoyeur de solutions pour développer la compétitivité des pays. Avec PISA, l'OCDE place en particulier, au cœur des discussions de politique éducative du principe de l'évaluation des résultats de l'éducation, résumée par le concept de « performance » utilisé par les pays anglo-saxons adeptes du Nouveau Management Public.

Elle se sert également de ses caractéristiques intrinsèques qui lui confèrent un « soft power » fort sur ses pays membres. L'organisation est en effet une référence internationale pour ce qui est des problématiques économiques, même si, on l'a dit, son expertise en la matière est mise à mal par le contexte de crise et la distanciation de certains pays de la doctrine qu'elle préconise. Mais elle demeure une organisation au pouvoir symbolique certain, du moins sur les pays européens, puisqu'elle a été le bras financier de la reconstruction des pays européens à partir des années 1960. De même, la stratégie du consensus qu'elle prône entre la myriade d'acteurs impliqués dans le processus PISA, à sa création comme à chaque nouveau cycle, participe à renforcer la légitimité des résultats de PISA et donc la crédibilité et le poids de l'OCDE comme experte des politiques éducatives. L'implication des administrations nationales, sur les plans aussi bien politique que technique, crée un sentiment d'appartenance à une communauté d'experts dont on verra qu'elle engendre un sentiment parallèle d'obligation de contribution.

L'OCDE se pose en avant garde avec PISA et se définit un rôle à la fois d'inspirateur pour l'évaluation de la politique éducative et de coordonnateur

²¹⁶ Vincent Gayon, « Le crédit vacillant de l'expert », *Cultures & Conflits*, 75, automne 2009, p.55.

stratégique et technique. Par cette position officiellement en recul de l'organisation concrète de chaque cycle, elle se distancie de l'analyse des résultats de l'éducation. Elle donne à voir à chaque pays participant, par le biais d'indicateurs unanimement acceptés par l'ensemble d'entre eux, une série de données statistiques individuelles au milieu de celles des autres, problématisés mais sans réaliser d'analyse des explications ni des implications possibles en termes de politiques éducatives. D'une certaine manière, un peu à la façon de l'œil du Persan, l'OCDE se décrit comme donnant à chacun une représentation factuelle des résultats de l'éducation, neutre du point de vue de l'analyse, mais qui incite chaque pays à se comparer aux autres participants. La neutralité de l'analyse n'est finalement que formelle, puisqu'elle transparait à travers le choix des indicateurs faits au moment de la publication principale et de ceux égrenés au fil des publications thématiques intermédiaires qui paraissent plusieurs fois par an et qui se focalisent sur certaines problématiques jugées prégnantes dans les débats de politique éducative du moment.

Enfin, l'OCDE se crée un rôle multifacette, à la fois d'évaluateur des politiques éducatives, de décideur donnant les orientations stratégiques, et d'expert technique à la pointe des méthodes statistiques appliquées aux politiques publiques. Nous verrons que ce positionnement assez inédit engendrera des frictions dans le champ de l'expertise nationale de l'évaluation des politiques éducatives françaises, et nécessitera un repositionnement aussi bien des évaluateurs que des décideurs politiques français vis-à-vis des agents de l'OCDE.

Partie 2 : L'irruption de nouveaux agents « experts » de l'évaluation des politiques éducatives : un repositionnement en deux temps des évaluateurs nationaux

Cette seconde partie interroge la façon dont l'arrivée de PISA questionne la fonction et la posture des évaluateurs nationaux et oblige les agents de la DEPP à reconfigurer leur champ d'action. Nous nous intéresserons tout particulièrement aux « évaluateurs de terrain », ceux chargés au sein de l'administration centrale de la mise en œuvre des évaluations nationales et du traitement des résultats, et qui alors sont chargés d'être les interlocuteurs techniques de l'équipe PISA-OCDE. Ils sont nos acteurs « pertinents », ceux que l'on regarde pour saisir notre objet de recherche, ici l'évaluation de la politique éducative menée par les évaluateurs français. Nous nous affranchissons volontairement de l'analyse des idéologies politiques qui sous-tendent les décisions prises par les décideurs publics, et regardons l'effet de la mouvance des rôles perçus sur la façon d'exercer une mission d'intérêt public. Nous considérons nos évaluateurs de terrain comme un groupe d'individus relativement homogène du point-de-vue de ce qui nous intéresse ici. Finalement, au-delà des individualités se crée une réponse commune à un questionnement mis en commun sur la place des agents de la DEPP dans le processus d'évaluation de la politique éducative française et sur leur identité d' « expert ».

Tout juste attribués d'une fonction d'évaluation des résultats des élèves désormais incontournable compte tenu de l'entrée franche du système français dans la vague du Nouveau Management Public, les évaluateurs de la DEPP doivent composer avec un inévitable travail commun avec l'équipe PISA-OCDE. Pilotes des évaluations des résultats des élèves françaises, ils se voient attribuer aussi bien par leur hiérarchie que par les équipes de PISA-OCDE une mission de « messenger » de la part de l'administration du ministère et de « répondant » aux demandes des équipes de l'OCDE et du consortium chargé de gérer la passation

des tests. Nous verrons comment la perception qu'ils ont de ce nouveau rôle influe sur la façon qu'ils ont de faire de l'évaluation. Pour cela, nous regarderons tout d'abord le rôle qu'ils se voient attribuer à l'orée du XXI^e siècle, à la fois par la nouvelle organisation de la gestion publique au moment paradigmatique de la mise en place de la LOLF, et à la fois par la direction de l'éducation de l'OCDE au moment où le pilotage des évaluations de PISA est pensé. Les parcours des agents et l'image qu'ils ont de la « fonction » qu'ils occupent sont des éléments de compréhension de la façon dont ils se positionnent. (Chapitre 3).

Ce rôle exécutif les amène d'abord à construire un cadre d'analyse critique du programme d'évaluation international. L'arrivée de PISA dans la boîte à outils des évaluations disponibles en France dans lesquelles les évaluateurs français sont impliqués au premier rang pourrait permettre aux évaluateurs comme aux décideurs d'avoir de nouvelles données de pilotage de l'action publique en éducation. Finalement ce n'est pas cet usage qui est privilégié au départ. Les résultats PISA, malgré l'investissement qu'ils ont nécessité, seront réduits à l'utilité minimum et serviront ensuite de tremplin aux évaluations « maisons » construites par la DEPP – qui communiquera sur les proximités méthodologiques entre ses nouvelles évaluations et PISA pour les valoriser auprès des décideurs politiques (chapitre 4).

Chapitre 3 : Évaluateurs nationaux et internationaux : une même mission à partager ?

L'évaluation des politiques éducatives se retrouve au cœur d'un double paradigme en France à la fin des années 1990 et au début des années 2000. Plusieurs événements concomitants en dessinent le rôle théorique. La gestion de la dépense publique est d'abord totalement remaniée par la Loi Organique relative aux Lois de Finances qui donnent à la statistique et au pilotage par indicateur un rôle central dans le pilotage de l'action publique. Dans le même temps, le paysage des évaluations des résultats de l'éducation s'enrichit d'un nouveau venu d'envergure, l'évaluation PISA, de dimension mondiale et pilotée depuis le siège de l'OCDE situé à Paris, dans laquelle l'administration française est engagée par l'intermédiaire de la DEPP.

Nous verrons dans un premier temps que semble se construire un cadre propice à la montée en puissance du rôle des évaluateurs de la DEPP, disposant de toutes les caractéristiques à la fois technique – leur maîtrise statistique et d'élaboration d'indicateurs – et administrative – leur statut de direction de l'évaluation au sein du ministère de l'Éducation nationale – pour être les interlocuteurs privilégiés des décideurs publics au moment d'élaborer de nouvelles réformes du système éducatif souhaitables au vu des indicateurs des résultats des élèves. Cela ne se matérialise pourtant ni par une reconnaissance publique de l'utilité de leur travail évaluatif pour piloter l'action publique ni par une valorisation de leur expertise, mise en concurrence avec celle des évaluateurs de l'équipe PISA-OCDE. Cette mise en concurrence n'est pas, on le verra, frontale et explicite. Elle repose sur la sollicitation d'acteurs plus que d'autres par les décideurs publics, sur la plus ou moins grande latitude laissée par l'équipe

PISA-OCDE dans la réalisation des évaluations internationale, et sur les écarts perçus par les agents de la DEPP entre le rôle qui leur était théoriquement attribué et le rôle qu'ils estiment avoir en réalité tout au long du processus de mise en œuvre des cycles de PISA.

Cette dissonance entre le rôle perçu et le rôle théorique amène à des réactions de réaffirmation des compétences et de la légitimité des agents de la DEPP en matière d'évaluation. Nous verrons dans un second temps que ceux qui auraient dû être, par leur fonction au sein de l'administration centrale, les experts de l'évaluation, peinent à faire valoir leur statut lorsqu'il s'agit d'intervenir au nom de la France auprès de l'OCDE pour demander des modifications jugées indispensables pour que l'évaluation soit la plus adaptée au système français possible et donc la plus représentative possible des résultats de la politique éducative menée. De la même façon, alors qu'ils sont posés par l'OCDE en interlocuteurs privilégiés dans la mise en œuvre de PISA, et que leur pouvoir d'influence sur le contenu et l'organisation de l'évaluation semble érigé en principe d'action fondamental par les pilotes de PISA au sein de l'organisation internationale, les échanges se font finalement davantage de l'OCDE vers la DEPP que l'inverse. Le pouvoir d'expertise des agents de la DEPP en matière d'évaluation statistique des résultats des élèves s'en trouve repositionné à une échelle française, interne, qui n'est pas celle attendue par eux.

Une réflexion centrale sera menée autour de la notion d'expertise et la façon dont elle est mobilisée dans la conception et la tenue des évaluations statistiques appliquées aux politiques éducatives. En quel sens les acteurs de l'évaluation, notamment de l'évaluation standardisée, sont-ils des experts ? Assiste-t-on avec l'arrivée de PISA à une concurrence d'experts sur un même terrain, ou influent-ils sur des champs différents, par notamment la différence de l'échelle de leurs études, des principes mêmes qui les régissent, des compétences et de l'organisation des équipes évaluatrices? Au-delà de la recherche d'une définition exhaustive de l'expertise, nous analyserons quels experts les évaluateurs nationaux et internationaux pensent-être. Et au-delà de la perception

de chaque individu sur sa propre expertise, nous observerons les caractéristiques de l'expertise de chaque groupe d'évaluateurs mis en concurrence, pour identifier les dimensions mises en avant par chacun pour asseoir son rôle d'aide à la décision publique. Ces dimensions de l'identité d'évaluateur permettront dans la suite de cette thèse d'apporter des éléments explicatifs à la façon dont les résultats PISA ont été reçus par ceux chargés d'en faire une analyse plus poussée et opérationnelle pour accompagner la décision publique en matière de politique éducative. Nous proposerons ici une approche sociologique des acteurs²¹⁷ de l'expertise et de l'évaluation en analysant les effets des formations et des trajectoires professionnelles de ceux-ci sur leur manière à la fois d'évaluer et de concevoir le rôle de l'évaluation. Une telle approche permet de saisir les accointances liées à la fréquentation de certains forums d'échange ou de lieux de discussions et de prise de décision. Cela permet aussi de saisir en quoi les perceptions des acteurs de l'expertise qu'ils allouent à leurs interlocuteurs influent sur la capacité à imposer leurs prises de position, et notamment leurs prises de position techniques.

1) Les évaluateurs nationaux au cœur de nouveaux paradigmes de l'évaluation de l'action éducative

L'OCDE, qui se conçoit à travers son équipe PISA comme un nouveau venu dans le monde des évaluateurs internationaux et veut assurer de son expertise en matière d'évaluation des politiques éducatives, crée sur mesure une évaluation à la fin de la décennie 1990 qui fait intervenir largement les évaluateurs nationaux français, dont les fonctions et les liens avec le politique varient au gré des changements ministériels. A cette période, les évaluateurs nationaux impliqués dans PISA, les agents de la DEPP, sont en effet encore en train d'imaginer leurs marges de manœuvre pour jouer un rôle dans la conception

²¹⁷ Philippe Braud, « Chapitre premier. L'étude des acteurs », Philippe Braud éd., *La science politique*. Presses Universitaires de France, 2017.

de la politique éducative. Et ce travail de positionnement se déroule en même temps qu'un double paradigme qui redistribue les cartes de l'évaluation des politiques éducatives : au niveau international d'abord on l'a dit, les évaluateurs doivent composer avec leur nouveau rôle d'interlocuteur privilégié mais sans réel pouvoir de l'OCDE dans le bon déroulement des évaluations PISA ; au niveau national ensuite, la LOLF se met en place, qui met le principe d'évaluation au cœur du pilotage de l'action publique sans qu'un rôle ne soit pour autant attribué à la DEPP qui soit à la hauteur de ses missions. C'est ce double écart entre rôle théorique et rôle vécu qui a un effet sur leur façon de se positionner par rapport à ceux qui deviennent alors des concurrents, et qui se traduira, on le verra, par une remise en question forte de l'évaluation internationale par rapport aux travaux d'évaluation nationaux.

a) Les indicateurs de la LOLF : pilotes de la construction des « possibles » des politiques publiques

La LOLF a signé l'entrée résolue de la France dans l'ère du Nouveau Management Public. Mise en débat au parlement et votée en 2001, elle a nécessité plusieurs années de mise en œuvre au vu de l'ampleur du chantier de réorganisation de l'ensemble du budget de l'Etat. Son déploiement sera effectif dans l'ensemble des secteurs dans lesquels intervient l'Etat à l'occasion des débats budgétaires pour l'année 2006.

Sur le fond, la LOLF a deux objectifs : donner un rôle de contrôle plus important aux parlementaires pour regarder la dépense publique, et entrer dans l'ère de la « modernisation de la gestion publique »²¹⁸ en passant d'une logique de moyens à une logique de résultat. Sur le premier objectif, il est à noter que la loi

²¹⁸ Damien Catteau, *La LOLF et la modernisation de la gestion publique : la performance, fondement d'un droit public financier rénové*, thèse pour le doctorat de droit public financier sous la direction de Xavier Vandendriessche, soutenue à l'Université de Lille 2 en 2005.

organique a été lancée à l'initiative de deux parlementaires convaincus de la nécessité de refondre la gestion publique française pour en améliorer le contrôle et l'efficacité : Alain Lambert, président de la commission des finances du Sénat, et Didier Migaud, rapporteur général de la commission des finances, de l'économie générale et du plan de l'Assemblée Nationale et futur Premier président de la Cour des Comptes qu'il rejoindra en 2010. Cette implication plus grande des parlementaires dans l'évaluation des politiques publiques nous intéressera dans la troisième partie de cette thèse lorsque nous analyserons comment PISA et les travaux de la DEPP seront mobilisés lors des débats législatifs sur les grandes lois éducatives de la période, en 2005 et 2013.

Sur le second objectif, au-delà de l'immense travail de reconstitution des règles budgétaires qui a été entrepris sur plusieurs années à partir du vote de la loi organique, la place des indicateurs décrivant à la fois l'action publique passée et l'action publique souhaitée à l'avenir est devenue prépondérante. Il s'est agi de concevoir les outils descriptifs de l'action publique dans toutes ses dimensions, donnant à voir à la fois ses réalisations sur l'année passée et les ambitions et les moyens alloués pour y parvenir sur l'année suivante. L'évaluation est réellement placée au cœur de la réflexion sur la constitution de la LOLF, et donc au cœur de la réflexion sur le devenir de l'Etat. Cette grande réorganisation a eu des effets bien plus larges que la seule refonte de la présentation des différents budgets de l'Etat. Nous ne reprendrons pas ici l'ensemble des implications de la LOLF sur l'organisation des comptes publics ni ce que signifie le passage à une représentation de l'action publique par des « programmes » de politiques publiques et des « objectifs » sur la façon dont elle est conçue par les décideurs. Ces questionnements ont fait l'objet d'abondants travaux²¹⁹ qu'il serait bien trop long de résumer ici et qui ne concernent pas le cœur de notre propos. En revanche, il faut noter le nouveau rôle de contrôle de l'utilisation des dépenses publiques pour chaque politique publique attribué aux indicateurs statistiques, et

²¹⁹ Voir par exemple Jean-René Brunetière, « Les indicateurs de la loi organique relative aux lois de finances (LOLF): une occasion de débat démocratique? », *Revue française d'administration publique*, Ecole nationale d'administration, vol. 1 n°117, 2006.

faisant intervenir les organes chargés de la statistique publique. Compte tenu de la nécessité d'avoir chaque année des données statistiques faisant état de la réalisation ou non des objectifs alloués l'année précédente, on imaginerait aisément que le pouvoir des agents chargés de concevoir et de faire tourner ses indicateurs serait particulièrement fort car il aurait un effet immédiat sur la façon d'envisager une question publique. Les indicateurs seraient alors chargés de décrire chaque rouage utilisé pour atteindre un objectif donné, déterminé en amont chaque année au vu des résultats des objectifs passés. Ils sont 1350 à supposer mesurer les résultats de l'action de l'Etat, éclairer les choix budgétaires du Parlement, juger l'action des responsables et orienter l'action de l'Etat²²⁰.

L'OCDE n'est jamais bien loin lorsqu'est racontée l'histoire de la LOLF par les gouvernements qui se sont succédés depuis sa mise en œuvre. Elle est résumée ainsi sur les supports de communication gouvernementaux :

«Avant tout, la loi organique relatives aux lois de finances de 2001 (LOLF) a pu être élaborée grâce à l'existence d'un consensus politique exceptionnel sur la nécessité de réformer l'ordonnance de 1959.

Ce consensus s'appuyait d'abord sur les travaux de la Cour des comptes et du Parlement portant sur l'utilisation de l'argent public et la transparence en matière budgétaire.

Les Travaux de l'Organisation de coopération et de développement économiques (OCDE) promouvant une nouvelle gestion publique ont aussi guidé l'élaboration de la LOLF. Ils ont démontré l'intérêt pour les États d'adopter une démarche de budgétisation orientée vers la performance, décrivant les politiques publiques, leur fixant des objectifs, laissant aux gestionnaires une grande latitude dans l'utilisation des crédits, avec pour contrepartie un système de contrôle plus rigoureux.

Enfin, la LOLF s'est également inspirée des modernisations des procédures budgétaires et des méthodes comptables menées par de

²²⁰ Jean-René Brunetière, *art. cit.*, Michel Lascombe, et Xavier Vandendriessche. « La loi organique relative aux lois de finances (LOLF) et le contrôle des finances publiques », *Revue française d'administration publique*, vol. n°117, n° 1, 2006. Voir également plus largement sur la LOLF et le Nouveau Management Public Philippe Bezes, *Réinventer l'Etat, les réformes de l'administration française (1962-2008)*, Paris, PUF, 2009.

nombreux États depuis les années 1980. »²²¹

Les mots clés sont présents dans ces quelques lignes, et expriment tout l'esprit de la LOLF, dans la lignée des doctrines économiques de l'OCDE. Il est intéressant de noter l'utilisation de plusieurs concepts qui sont censés, aux yeux du rédacteur de cette présentation s'exprimant au nom du gouvernement français, démontrer le bienfondé de cette réforme et les avantages qu'elle procure par rapport à l'ancien modèle de gestion publique. Il invoque trois types d'arguments. Le premier est un argument de légitimité démocratique. La LOLF aurait été mise en œuvre avec l'accord de tous les partis, son intérêt étant reconnu de façon transpartisane elle se positionne au-dessus du débat politique, au niveau de l'intérêt général.

Le deuxième argument porte sur la justesse des principes qui la sous-tendent, démontrée par les études économiques de l'OCDE, tenant lieu ici d'autorité validant par sa simple identité l'argument donné. Les études de l'organisation internationale, menées sur les systèmes de gestion publique de ses pays membres, auraient démontré la pertinence d'un modèle de performance, de résultat donc. Le rôle des indicateurs de résultat tire ici toute son importance. Les performances réalisées nécessitent d'être évaluées régulièrement, et, puisqu'une latitude est laissée aux gestionnaires dans l'atteinte des objectifs fixés, ils impliquent de travailler sur ce que l'on souhaite évaluer. La démarche d'évaluation n'est pas seulement celle des dépenses effectuées, des moyens mobilisés qu'ils soient budgétaires ou humains, mais bien de la réalisation ou non des objectifs pré-déterminés, qui témoignent des orientations souhaitées de l'action publique.

Le troisième argument, qui en dérive presque, est celui de l'exemple et du retour d'expérience donné par « plusieurs » pays depuis « plusieurs » décennies. Ce recul permettant de constater selon l'auteur de la description la validité à la fois dans le temps et dans plusieurs configurations nationales différentes est alors

²²¹ <http://www.vie-publique.fr/fiches/21907-comment-la-lolf-a-t-elle-ete-preparee-et-adoptee>, consulté le 12 juin 2018.

censé confirmer la cohérence de la loi dans le cas du système français. Cette vision homogène de l'existence de « meilleures pratiques » adaptées à différents contextes est tout à fait dans la lignée des travaux de l'OCDE dont l'objectif est bien de fournir une réponse collective à des problématiques de politiques publiques rencontrées par ses membres²²².

Cette simple définition en ligne sur un site gouvernementale témoigne bien de deux choses. D'abord, elle rend visible l'importance croissante à l'époque de la LOLF des fonctions d'évaluation et de construction d'indicateurs pertinents dans le pilotage de l'action publique, y compris en éducation. Manifestation claire d'une « obsession du contrôle ²²³», la LOLF systématise ce dernier en structurant une action administrative autour d'outils quantitatifs faisant intervenir au premier plan les méthodes statistiques. Elle montre ensuite la puissance de l'OCDE dans la justification de la mise en œuvre de politiques nationales. Si l'on revient vers nos évaluateurs nationaux, on voit donc qu'avec la mise en place de la LOLF, les évaluateurs de la DEPP deviennent à la fois les responsables des évaluations des résultats des élèves et du système d'enseignement, mais aussi les détenteurs de l'outil de gestion désormais principal de la comptabilité nationale. L'outil de pilotage de la politique éducative devient un outil de gestion comptable. Ils tiennent là un rôle théorique central dans ce début de millénaire pour le pilotage de la politique éducative, puisqu'ils sont non seulement organe évaluateur interne au Ministère mais aussi interlocuteur privilégié, pour la France, de l'OCDE dans la mise en œuvre de ce qui deviendra l'évaluation internationale la plus visible jamais menée, PISA.

²²² Cet objectif de fournir une réponse « globale » a été notamment bien rappelé par Angel Gurría, secrétaire général de l'OCDE de 2006 à 2021, dans son discours célébrant les 50 ans de l'OCDE le 24 mai 2011 au siège de l'OCDE à Paris: « nous sommes à pour forger une réponse collective » dit-il au premier point de son discours.

<https://www.oecd.org/fr/apropos/secretairegeneral/betterpoliciesforbetterlives.htm>

²²³ Philippe Bezes, « Le renouveau du contrôle des bureaucraties. L'impact du New Public Management », *Informations sociales*, vol. 126, n°6, 2005 ; plus largement sur le Nouveau Management Public voir son ouvrage *Réinventer l'État, les réformes de l'administration française (1962-2008)*, Paris, PUF, 2009.

b) Les travaux de la DEPP : un outil de la nouvelle gestion publique ?

Instrument de gestion « *doté d'une technologie invisible qui structure les représentations* »²²⁴, la LOLF induit donc, par ses principes de contrôles dans l'intérêt collectif et emprunts d'une modernité recherchée par nombre d'Etats, une hiérarchie implicite des évaluateurs. Ceux qui ont la main sur ses indicateurs considèrent contribuer à rendre la gestion publique française exemplaire et transparente. Ils contribuent en creux à déterminer les objectifs de l'action publique car ils définissent ce qui est du ressort de l'évaluable ou non, et donc ce qui est du ressort des moyens de l'Etat ou non. Ce sur quoi l'Etat ne peut pas agir ne peut finalement pas lui être imputé, et tout l'enjeu de la construction des indicateurs est bien alors de montrer scientifiquement la façon dont l'Etat est intervenu pour atteindre les objectifs qu'il s'est fixés. Ils montrent les « *résultats de l'action de l'Etat* », qui, s'ils ne sont explicités réellement nulle part, dans aucun texte législatif, sont pourtant analysés régulièrement par les rapports écrits par les assemblées parlementaires sur de multiples sujets d'intérêts pour les élus en mandat. Et la définition retenue se fait bien souvent par exclusion de ce que les résultats de l'action ne sont pas, à la manière de celle énoncée par Jean-René Brunetière: « *Les résultats qualitatifs de l'action de l'Etat, au-delà de la production de l'administration (le nombre d'actions menées dans un domaine, par exemple le nombre de contrôles d'alcoolémie sur les routes), ne sont pas ce sur quoi il n'a aucune influence exclusive ou déterminante.* »²²⁵

On voit bien là l'enjeu de construction des indicateurs permettant d'évaluer la politique éducative française, sur laquelle l'Etat a historiquement un pouvoir fort et de laquelle il fait une de ses prérogatives phares. Il a bien de l'influence déterminante sur la politique éducative, même si évidemment toute la difficulté réside dans la multiplicité des sources « éducatives » des petits Français, que ce soit dans la sphère familiale, amicale ou tout simplement dans les situations de

²²⁴ Ariel Eggrickx, « Réflexion critique sur l'adoption d'outils de gestion par mimétisme : le cas de la LOLF », *Management et Avenir*, Management Prospective Ed., 4 n°54, 2012, p.62.

²²⁵ Jean-René Brunetière, *art. cit.*

la vie quotidienne, et qui la rendent non-exclusive à l'action de l'Education nationale.

Dans ce cadre, et bien qu'ils soient en situation monopolistique sur les évaluations quantitatives du système éducatif français²²⁶, les évaluateurs de la DEPP n'ont pas été sollicités pour construire les nouveaux indicateurs de la LOLF. En effet miroir, la LOLF n'a pas non plus eu d'impact réel sur les évaluations menées par la DEPP en n'étant pas à l'origine par exemple du développement de nouvelles méthodologies d'évaluation. Ce moment où la DEPP aurait pu devenir l'interlocuteur privilégié du gouvernement pour réfléchir à la façon de représenter le contrôle des objectifs assignés à l'Education Nationale n'a pas eu lieu. Finalement, peu de nouveaux indicateurs voient le jour pour le budget de l'Education Nationale, qui reste, dans les premières années de la LOLF, représenté par une compilation de données déjà produites par les organismes de statistiques d'Etat comme l'INSEE et de données très descriptives issues d'évaluations de l'état du système éducatif menées de longue date par la DEPP²²⁷. L'absence de sollicitation pour créer de nouvelles dimensions évaluatives et pour réfléchir à la façon la plus juste d'évaluer l'atteinte des objectifs qualitatifs fixés par l'Etat en matière d'éducation est vécue par beaucoup d'agents comme une mise à l'écart. Certains répondront que les indicateurs finalement ne permettent pas de représenter de façon fidèle les orientations de politique éducative, et de politique publique en général, décidée par le gouvernement. Les indicateurs, par leur nombre limité bien que conséquent, ne donneraient qu'une vision partielle de la direction prise par l'Etat une année donnée et se rapporteraient davantage à des priorités qu'à réellement l'ensemble des objectifs fixés²²⁸. Toutefois, symboliquement et au-delà des considérations de représentativité du réel, cette absence de sollicitation vient d'ajouter à un contexte

²²⁶ Xavier Pons, « Figures de l'expert international français : l'exemple des évaluations internationales des acquis (1958-2008), « Carrefours de l'éducation », n°37, Armand Colin, 2014.

²²⁷ Entretien avec Agathe Cagé, Conseillère au cabinet du ministre de l'Education Nationale en charge du second degré, des programmes et de l'évaluation des élèves, octobre 2016.

²²⁸ Jean René Brunetière, « Les indicateurs de la loi organique relative aux lois de finance (Lolf) : une occasion de débat démocratique ? » *Revue française d'administration publique*, n°117, 2006.

déjà déroutant pour les évaluateurs de la DEPP, alors que cette dernière se positionnait il y a peu comme « *le champion du développement nécessaire d'une culture de l'évaluation* » ». ²²⁹

A l'époque où la LOLF est conçue et mise en place, la DEP est devenue DPD on s'en souvient. La crise d'identité de la direction avait été partagée auprès de parlementaires des groupes de gauche, et jugée suffisamment légitime pour faire l'objet d'une question au gouvernement, en direct en séance au Sénat. La sénatrice Nicole Borvo, sénatrice communiste de Paris, avait ainsi alerté en janvier 2000 le ministère sur « *la situation préoccupante que traverse la direction de la programmation et du développement.* » ²³⁰ Elle insistait sur la remontée de la part des équipes en place de « *dysfonctionnements lourds qui contribuent à une détérioration continue de l'audience et des conditions de travail* ». Et elle ajoutait : « *Nombreux sont les personnels qui, au sein de la DPD, pensent que celle-ci est un chantier en panne, sans mission claire, sans projet précis ni pilotage sérieux.* » Quel paradoxe : à l'époque où le gouvernement travaille à la LOLF, en même temps que le ministère de l'éducation, par la voix de la DEPP, investit des moyens humains et financiers conséquents dans sa participation à PISA, les budgets alloués à la direction ne seraient pas suffisants pour leur permettre de mener à bien leurs travaux. Une fois posé le constat de ce sentiment de déconsidération en interne à la DEPP qui se cristallise autour des budgets alloués, la question orale et la réponse qui y est apportée par Ségolène Royal, alors ministre déléguée chargée de l'Enseignement scolaire, laissent à voir la difficulté qu'éprouvait la direction elle-même, le ministère, et les parlementaires, à la positionner unanimement et dans le champ de l'évaluation de la politique éducative.

Le « chantier en panne » dépeint par Nicole Borvo concerne selon elle à la fois les moyens et les objectifs alloués aux missions de la direction. La DEPP a besoin du budget nécessaire pour mener à bien ses publications et les enquêtes dont elle a la charge, *Géographie de l'école* et l'enquête sur la situation des jeunes

²²⁹ Xavier Pons, *art. cit.* reprenant une formule de Claude Thélot en 1994, p. 72.

²³⁰ Question orale n°0692S de Mme Nicole Borvo (Paris, CRC) publiée dans le JO du Sénat du 19/01/2000, p. 81.

lycéens dans la vie active. Elle a sa place au Salon de l'Education et ses publications doivent être mises en valeur auprès des acteurs de terrain, enseignants, pédagogues. Elle se place également en pourvoyeuse de « *statistiques de la recherches* » très spécifiquement auprès des « *instances internationales comme l'organisation de coopération et de développement économique* ». La sénatrice ajoute la dimension d'éclairage de l'action publique et de rendu compte auprès des citoyens de l'utilisation des deniers publics :

« Il est également essentiel d'assurer la transparence, l'accès au public le plus large à une information indispensable, vous en conviendrez, pour évaluer les performances du système éducatif. (...) Les services de la DPD, qui effectuent ces recherches, ne devraient-ils pas voir leurs moyens renforcés, en tout cas disposer des moyens adéquats pour maintenir, voire pour renforcer, la production de données, travaux et études statistiques, nécessaires à l'éclairage de la politique de l'Etat comme à l'information du plus large public possible dans un domaine aussi sensible ? »

La ministre dans sa réponse ajoute à tout cela la fonction, centrale selon elle, de suivi de l'application des réformes menées par le ministère et en réaffirme la place supposée primordiale dans la dynamique de modernisation de l'Etat :

« La direction de la programmation et du développement est donc l'un des éléments importants de la modernisation du ministère de l'éducation nationale, de la recherche et de la technologie. Je le répète : l'activité de cette direction du ministère ne se réduit pas aux publications, même si c'est effectivement son activité la plus visible. En effet, cette direction a été chargée par les ministres concernés d'évaluer et de suivre l'application des réformes en cours et de l'ensemble des décisions ministérielles. Il s'agit là d'une mission fondamentale, qui donne lieu à un travail intense. »²³¹

Et là encore le rôle de la DEPP comme lien avec l'OCDE est réaffirmé et valorisé. L'OCDE n'est pas un interlocuteur final, les données produites pour l'organisation internationales – et à cette époque PISA connaît son premier cycle

²³¹ Réponse de Ségolène Royal à la question de Nicole Borvo, en séance au Sénat.

– ont bien pour objectif, selon la ministre, d’être complétées et enrichies pour être employées à des fins de pilotage de l’action publique en éducation:

«Les statistiques sur la recherche élaborées par la direction de la programmation et du développement, qui sont utilisées par les organisations internationales, comme l’OCDE, doivent être effectivement enrichies afin de pouvoir aussi fournir les outils d’aide à la décision permettant de répondre aux demandes de ceux qui sont chargés de piloter le système éducatif. »

Ces déclarations officielles de la ministre devant la représentation nationale, qui place la DEPP au centre du fonctionnement du ministère de l’Education nationale sont bien peu en rapport avec les sollicitations réelles de la direction pour la mise en place de la LOLF.

Pour les équipes en place de la DEPP, censée dans les textes et dans les mots des décideurs publics mesurer les résultats des élèves et donner des outils, des données, pour penser le pilotage de l’action de l’Etat en matière d’éducation, il est difficile de voir que leurs évaluations ne sont pas reprises réellement dans les indicateurs éducatifs. L’évaluation CEDRE notamment, qu’ils mènent depuis 2001 et pour laquelle une méthodologie nouvelle inspirée à plusieurs égards de PISA a été employée²³², n’est pas retenue. Malgré l’investissement des équipes de la DEPP et les intentions affichées du ministre Jack Lang²³³ quelques années plus tôt de créer de nouvelles évaluations internes à la DEPP pour mesurer les compétences globales des élèves, CEDRE ne convient pas aux critères requis par le cabinet du ministre pour les indicateurs de la LOLF.

« Jack Lang avait dit par l’intermédiaire de son directeur de cabinet adjoint qu’il fallait évaluer les compétences générales, avec l’idée que ça pourrait servir à la LOLF à venir. Le conseiller éducation en 2005 n’a pas retenu CEDRE, parce que les disciplines étaient à chaque fois

²³² Nous reviendrons plus en détail sur les proximités entre PISA et CEDRE et les stratégies d’acteurs qui ont accompagné la création de CEDRE par la DEPP au chapitre 4.

²³³ Jack Lang était ministre de l’Education nationale de mars 2000 à mai 2002.

différentes. »²³⁴

Si cela est ressenti par les agents au moment de la construction des indicateurs de la LOLF, les documents distribués aux parlementaires au moment du vote des textes budgétaires, leur fournissant tous les résultats de chaque indicateur sélectionné pour représenter les actions menées durant l'année et les actions prévues l'année suivante, sont parlants. L'examen des « bleus » budgétaires, comme sont appelés ces documents formalisés par des livrets à la couverture bleue, permet notamment de regarder la place accordée aux données fournies par la DEPP à deux égards. D'abord à travers leur usage pour rendre compte de l'action de l'Etat au moment de l'examen du budget. Ensuite à travers les moyens alloués à la direction pour réaliser ses missions. L'organisation du budget de l'Education nationale en programme est effective au moment de l'examen du projet de loi de finance pour 2006, en fin d'année 2005.

Si l'on regarde le bleu de la mission Enseignement Scolaire pour l'année 2006, les données fournies par la DEPP restent très descriptives de l'état du système, peu orientées vers des indicateurs qualitatifs comme les résultats des élèves. Pêle-mêle, ils donnent les chiffres des effectifs par niveau, du nombre d'enfants primo-arrivants accueillis dans le système scolaire, le nombre d'enseignants affectés au remplacement. Quelques indicateurs prévisionnels indiquent que les résultats des élèves en français et en mathématiques dans les écoles classées ZEP-REP et celles non ZEP-REP pourront être fournis les années suivantes pour évaluer l'objectif d'accroître la réussite des élèves en zone difficiles et des élèves à besoins éducatifs particuliers, mais ces indicateurs étaient déjà existants par le passé²³⁵.

Du point-de-vue des moyens accordés par l'Etat à l'évaluation et à la DEPP en particulier, rien ne montre une montée en puissance du rôle qui lui est donné. Les budgets attribués au programme évaluation sont plutôt stables en

²³⁴ Entretien avec Bruno Trosseille, Directeur du bureau des évaluations des élèves, DEPP, 14 avril 2016.

²³⁵ Bleu budgétaire, PLF 2006, Programme 104.

proportion de la totalité des budgets de soutien à l'action du ministère de l'éducation nationale, autour de 2,5%, alors qu'en parallèle de plus en plus de missions sont allouées à la Direction. Les bleus budgétaires des projets de loi précisent les objets de travail de la DEPP année après année. En 2006, la direction est mentionnée rapidement, sans détail des travaux évaluatifs, au milieu des acteurs de l'action n°2 « Evaluation et contrôle » :

« Les aspects pédagogiques et l'évaluation individuelle des enseignants relèvent de l'inspection générale de l'éducation nationale (IGEN) et des inspecteurs pédagogiques régionaux. La mission de contrôle, d'évaluation et de conseil de l'inspection générale de l'administration de l'éducation nationale et de la recherche (IGAENR) porte sur les aspects administratifs, financiers, comptables de l'organisation de toutes les structures, de statut public ou privé, recevant des fonds publics, Le Haut Conseil de l'évaluation de l'école (HCEE), autorité administrative indépendante, a pour mission de rendre l'information sur l'école la plus indépendante et transparente possible et de favoriser ainsi la qualité du débat public sur l'éducation. La direction de l'évaluation et de la prospective (DEP) définit et met en œuvre le dispositif d'évaluation du système éducatif et contribue à l'évaluation des politiques conduites par le ministère. »²³⁶

Quatre ans plus tard, coïncidant avec la sortie des résultats du troisième cycle PISA à l'hiver 2010, les évaluations sont cette fois détaillées une par une et témoignent de l'ampleur de la tâche allouée à la DEPP à moyens quasi constants :

« A ce titre, les services du ministère et le Haut Conseil de l'Education sont chargés :

- de réaliser des évaluations (PISA « lecture », bilans de fin d'école et de fin de collège dans le domaine de la maîtrise de la langue et des compétences générales, tests de mathématiques et de français, généralisation des projets européens « Apprendre à apprendre » et « maîtrise des compétences en langues vivantes »), d'études et d'enquêtes statistiques sur le monde éducatif (étude Surveylang sur les compétences linguistiques, panel élèves 1995, panel bacheliers 2008, participation à l'enquête emploi du temps 2009 de

²³⁶ PLF 2006, extrait du bleu budgétaire de la mission : Enseignement scolaire. Programme 214 : Soutien de la politique de l'éducation nationale, p. 7.

l'INSEE...);

- de développer, améliorer et procéder à la maintenance des systèmes informatiques décisionnels et des bases de données : SAS, ORACLE, Base Nationale des Identifiants Elèves... ;

- de produire les cahiers d'évaluation et publier un certain nombre d'ouvrages sur le système éducatif, parmi lesquels « L'état de l'école », « La géographie de l'école », les « Notes d'information », les « Références statistiques » et la revue « Education et formation » ;

- d'assurer l'impression et le routage de la plupart des productions. »²³⁷

On remarquera que les évaluations PISA sont citées en premier dans la liste, témoignant de la place primordiale qu'elles occupent déjà au bout de trois cycles dans le paysage des évaluations des résultats des élèves utilisées par les décideurs publics français.

Au lancement de PISA les agents de la DEPP se trouvent donc plongés dans un contexte très favorable à la montée en puissance de leur rôle au sein des évaluateurs de la politique éducative française. Pilote d'évaluations devenues centrales dans la conduite de la gestion de la politique publique, ils sont également nommés interlocuteurs principaux de l'OCDE dans le cadre de la mise en route de la plus grande évaluation internationale des résultats de l'éducation. Ils ont en théorie la confiance du ministère pour représenter la France et faire remonter lors des réunions avec leurs homologues et les équipes de PISA-OCDE les volontés de modification de certains items jugés inappropriés au système scolaire français ou les interrogations de méthode. Leur rôle de représentation se complète d'une fonction de médiation quasi diplomatique lorsque certains conflits ou certaines tensions émergent entre le ministère et l'OCDE. La DEPP est même laissée presque seule interlocutrice des équipes de l'OCDE en charge de PISA, y compris aux moments des changements de majorité présidentielle.

²³⁷ PLF 2010, mission « Enseignement scolaire », programme 214 Soutien à la politique du ministère de l'Éducation nationale, justification au premier euro.

Aux dires même des conseillers des cabinets ministériels, c'est à elle de rendre compte des orientations souhaitées par le Ministère, des problématiques techniques potentielles identifiées, avec une autonomie d'intervention et d'alerte forte. Une conseillère ayant travaillé auprès du ministre Vincent Peillon indiquait ainsi en entretien : « *C'est la DEPP qui est le lien avec l'OCDE sur PISA. Il n'y a pas d'intervention du ministre sur les tests, pas de projets politiques partagés.* »²³⁸

Pourtant, malgré ce rôle charnière théorique de la DEPP qui émerge au début des années 2000, ni son budget, ni l'utilisation publique de leurs travaux n'est à la hauteur des attentes en interne. L'investissement demandé par PISA est tel qu'ils sont amenés à plusieurs reprises à devoir délaissier un temps leurs évaluations « maison » pour se concentrer, à moyens humains constants, à l'évaluation de l'OCDE.

« On a quand même un assèchement des ressources financières de la DEPP pour participer à toutes ces études, au détriment du financement des études nationales. (...) Le regard international est important mais il entraîne un désengagement au niveau national, notamment financier et notamment sur les évaluations sur les programmes ».²³⁹

La mise en place de la LOLF pouvait impliquer pour la DEPP à la fois l'attribution d'une mission de création et de production régulière d'indicateurs centraux dans le pilotage de l'action éducative et à la fois l'entrée dans une période faste pour les évaluations statistiques mises au premier plan de la gestion publique. En parallèle, le lancement de PISA pour lequel ils sont nommés interlocuteurs pour le ministère de l'Education nationale tout en disposant d'une mission de conseil technique renforce leur statut d'évaluateur national à la légitimité affirmée après une période de remise en question de l'apport d'une direction d'évaluation interne au ministère. Cependant, en pratique, la

²³⁸ Entretien avec Agathe Cagé, Conseillère au cabinet du ministre de l'Education Nationale en charge du second degré, des programmes et de l'évaluation des élèves, octobre 2016.

²⁵⁰ Entretien avec Bruno Trosseille, Directeur du bureau des évaluations des élèves, DEPP, 14 avril 2016.

sollicitation de leur expertise pour la constitution des indicateurs de la LOLF reste réduite, et si la contribution à PISA est bien inscrite en première place de leurs missions, les moyens alloués ne s'en trouvent pas pour autant augmentés. Ce paradoxe entourant l'importance du rôle de la DEPP à cette période de renforcement de l'institutionnalisation de l'évaluation des politiques nationales en éducation constitue le contexte dans lequel s'organise le positionnement de ses agents, qu'ils font évoluer pour clarifier leur rôle d'expert de l'évaluation de l'action éducative française, notamment en regard de l'expertise allouée aux agents de l'OCDE en charge de PISA.

2) DEPP et OCDE: le « fonctionnaire savant » et le « fonctionnaire expert »

Les agents de la DEPP jouissent, selon le statut accordé à leur direction, d'une certaine liberté d'action pour imaginer de nouveaux outils évaluatifs et pour analyser finement les résultats obtenus dans les évaluations menées au niveau national. Ce rôle théorique n'est pas celui qui leur est attribué en pratique dans le cadre de leur travail sur PISA. Ils doivent, avec les équipes de l'OCDE, davantage répondre aux demandes du ministre que participer à une dynamique créatrice pourtant primordiale dans leur façon d'envisager leur rôle d'évaluateur. Une relation ressentie comme asymétrique du côté des évaluateurs français s'installe entre les équipes PISA d'un côté et les équipes de la DEPP travaillant sur PISA de l'autre, alors même que leurs trajectoires à la fois professionnelle et dans l'enseignement supérieur sont assez similaires et que la fonction de leurs postes, appartenant au fonctionnariat national pour les uns et international pour les autres, obéissent aux mêmes règles déontologiques. Au-delà de leurs statuts officiels et des rôles attribués à chacun par les documents fondateurs de PISA, un discours organisé se met en place dans chaque groupe pour définir leur expertise en fonction d'une série de caractéristiques donnant à leurs travaux une validité propre aux échelles de valeur internes à chaque organisation.

Afin d'étudier la façon dont le statut d'expert des évaluations de l'action éducative est revendiqué par nos évaluateurs, nationaux et internationaux, il faut d'abord s'interroger sur les contours de l'expertise dont il est question et en regarder les grands traits qui pourraient constituer des objets d'intérêt pour l'observation de nos évaluateurs nationaux et internationaux.

La définition de l'expertise n'est pas simple à formuler. L'expertise et les experts ont fait l'objet de nombreuses recherches depuis les années 1980, « à la mesure du succès public de l'appellation » en France²⁴⁰. Historiquement, l'expertise comme outil de la prise de décision a d'abord émergé dans le champ juridique, comme « *procédure par laquelle on confie à un ou plusieurs experts le soin de donner un avis sur les éléments d'un différend, quand ceux-ci présentent des aspects techniques.* »²⁴¹ La reconnaissance par le demandeur d'expertise d'une maîtrise technique chez l'expert, invité pour trancher, s'est progressivement étendue à d'autres questionnements techniques et/ou scientifiques et s'est ancrée dans le champ de la politique publique. Le lien entre expertise et administration est ancien, avec la constitution par l'Etat à la fin du XVIIIe siècle de grands corps issus d'établissements de formation considérés comme prestigieux (l'Ecole des Mines, l'Ecole des Ponts et Chaussées etc.), dont les anciens élèves étaient en charge d'éclairer l'administration²⁴² sur les conditions de réalisation de différents ouvrages publics.

Toutefois, la revendication de la nécessité de l'expertise dans le débat public pour des problématiques stratégiques d'action publique, est relativement récente et s'est invitée d'abord dans les prises de décisions politiques relevant de choix scientifiques dans la période d'après-guerre, impliquant de répondre à des situations d'incertitudes de plus en plus fréquentes.²⁴³ L'expertise est alors celle

²⁴⁰ Jean-Yves Trépos, *La sociologie de l'expertise*, PUF, Que-sais-je ?, 1996.

²⁴¹ Trésor de la langue française, « Expertise » [En ligne] : <http://atilf.atilf.fr/>. cité par Guillaume Calafat, *art.cit.*

²⁴² Renaud Bouvet, L'indépendance de l'expert, *Revue Juridique de l'Ouest*, Année 2015.

²⁴³ Laurence Dumoulin, Stéphane La Branche, Cécile Robert, Philippe Warin (dir.), *Le recours aux experts. Raisons et usages politiques*, Grenoble, Presses Universitaires de Grenoble, 2005 ; Daniel Benamouzig, Julien Besançon, « Administrer un monde incertain : les nouvelles bureaucraties techniques. Le cas des agences sanitaires en France », *Sociologie du travail*, 47 (3), 2005.

de scientifiques, médecins, biologistes, ingénieurs, appelés à apporter au décideur politique l'état des connaissances scientifiques à l'instant T de la prise de décision. Ils aident à fonder le choix des orientations prises - souvent dans des champs de santé publique ou environnemental - sur des savoirs faisant office de « preuves », objets d'un relatif consensus au sein de la communauté scientifique au moment où ils sont sollicités comme éléments d'expertise. La vitesse des prises de décision, en cas de crise sanitaire ou de catastrophe environnementale par exemple, dans un contexte de redevabilité croissante des décideurs envers leurs administrés et de transparence accrue des choix faits par les dirigeants, a amené par ailleurs ces derniers à construire un processus de prise de décision justifié avec autant de rigueur et d'information que le nécessite les responsabilités qui sont les leurs : celles d'être les garants de l'intérêt général par l'utilisation des deniers publics. Là intervient l'expert, qui, par sa capacité à identifier et à justifier les connaissances nécessaires à une prise de décision éclairée permet de tenir les délais d'une prise de décision qui serait à la fois responsable et efficace.²⁴⁴

L'expert en ce sens ne se contente pas d'apporter des éléments d'information, mais il doit aussi être capable de fournir au décideur les réponses à une problématique posée²⁴⁵. L'expertise engage alors son expert, qui fait le choix des éléments qu'il porte à la connaissance des décideurs qui le sollicitent, selon deux principes forts: indépendance et neutralité. L'indépendance est posée en condition indispensable à la validité de l'expertise produite.²⁴⁶ Le questionnement de l'indépendance de l'expert, central nous le verrons dans la définition et la reconnaissance de l'expertise des évaluateurs de la DEPP comme de l'OCDE, contribue à garantir le fait que la prise de décision sera fondée sur une donnée ou un avis objectif. Dans le cas de l'action publique, l'indépendance

²⁴⁴ Voir par exemple Jean-Yves Trépos, *La sociologie de l'expertise*, PUF, Que-sais-je ?, 1996 ; Cécile Robert, « Expertise et action publique », dans Olivier Borraz, Virginie Guiraudon (dir.), *Politiques publiques. 1, La France dans la gouvernance européenne*, Paris, Presses de Sciences Po, 2008.

²⁴⁵ Guy Lapostolle, *Du bon usage de l'expertise dans les politiques éducatives*, Politiques et management public, Vol. 24, n°1, 2006.

²⁴⁶ Christiane Restier-Melleray, « Experts et expertise scientifique. Le cas de la France », *Revue française de science politique*, 40-4, 1990.

de l'expert est alors d'autant plus nécessaire qu'elle permet de garantir que les décisions sont prises dans le respect de l'intérêt général et de la bonne gestion des deniers publics.

Au-delà de l'objectif d'une réponse rapide à une situation d'incertitude, la recherche d'une expertise indépendante et compatible avec la recherche de l'intérêt public entraîne une reconfiguration de l'action publique pour faire intervenir soit des experts privé²⁴⁷, soit institutionnalisant la mobilisation de l'expertise en l'intégrant aux pratiques administratives dans la lignée des principes du Nouveau Management Public²⁴⁸. Les deux ne sont d'ailleurs pas incompatibles, certains travaux mettant en évidence que les consultants ou évaluateurs privés sollicités en tant qu'experts par les décideurs publics jouent le rôle de passeurs de principes et d'outils de gestion du secteur privé vers la sphère publique²⁴⁹. Les évaluateurs deviennent experts lorsqu'ils sont plongés dans une « situation d'expertise », à la « rencontre d'une conjoncture problématique et d'un savoir spécialisé. »²⁵⁰ Ils mobilisent leurs savoirs pour répondre au problème, qui deviennent une expertise au sens pragmatique où l'entend Corinne Delmas dans la définition qu'elle propose en 2011 : sont expertises tous « savoirs mobilisés pour l'action »²⁵¹. Dans le cas de l'évaluation statistique qui nous intéresse ici, cela inclut donc à la fois les données produites sur demande d'un commanditaire et l'analyse contextualisée de données déjà existantes permettant d'apporter des éléments de conseils pour la prise de décision.

Plusieurs courants de recherche sur l'expertise en action publique ont ainsi émergé ces dernières décennies, selon plusieurs grandes approches dont nous retiendrons les suivantes pour étudier les manifestations et le repositionnement

²⁴⁷ Isabelle Berrebi-Hoffmann et Pierre Grémion, « A quoi servent les experts ? », *Cahiers internationaux de sociologie*, n°126, 2009.

²⁴⁸ Philippe Bezes, « Aux origines des politiques de réforme administrative sous la Ve République : la construction du « souci de soi de l'Etat » », *Revue Française d'administration publique*, 2, n°102, 2002.

²⁴⁹ Corinne Delmas, *Sociologie politique de l'expertise*, Paris, La Découverte, 2011.

²⁵⁰ Jérôme Péglise, Caroline Protais, Keltoume Larchet, Emmanuel Charrier (dir.), *Des chiffres, des maux et des lettres. Les experts judiciaires en économie, en psychiatrie et en traduction*, Paris, Armand Colin, coll. « Recherches », 2012.

²⁵¹ Corinne Delmas, *Sociologie politique de l'expertise*, Paris, La Découverte, 2011.

le cas échéant de l'expertise des évaluateurs nationaux de la DEPP par rapport à celle revendiquée par les agents de l'OCDE. Abordée sous l'angle des compétences détenues par les individus, la notion d'expertise interroge l'origine de leur acquisition, par l'expérience ou par la formation initiale ou continue. Elle peut provenir de l'expérience acquise aussi bien professionnellement que par un engagement citoyen, ou par des situations rencontrées dans le champ de la vie privée.²⁵² Abordée sous l'angle de la légitimité qu'elle suppose²⁵³, le statut d'expert n'est plus conféré par une formation ou une expérience particulière, mais par la reconnaissance de la valeur et l'expertise par son commanditaire et par tous ceux susceptibles de juger le processus aboutissant à une prise de décision. Il est donc relatif à l'organisation du secteur de l'action publique dans lequel il se place²⁵⁴, dans lequel s'articulent des rapports de pouvoir entre les acteurs.

Les compétences n'ont alors pas de valeur spécifique les unes par rapport aux autres dans l'absolu, mais deviennent source de légitimité d'un expert lorsqu'elles correspondent à celles valorisées par le commanditaire de l'expertise et acceptées par ceux qui seront les destinataires de la politique publique fondée sur l'expertise de l'expert. Dans le cas de la prise de décision en action publique, le ou les décideurs s'adressent à celui qu'ils jugent compétents pour répondre à notre questionnement. Celui qui désigne l'expert désigne dans le même temps les critères qui font expertise : elle a trait alors à la reconnaissance d'une institution ou d'un acteur reconnu par ses pairs²⁵⁵. En s'adressant à un individu pour son

²⁵²On citera notamment l'émergence de la figure du patient-expert dans le champ sanitaire et social et la valorisation par les professionnels de santé du savoir expérientiel dans l'élaboration de dispositifs d'accès aux soins et d'accompagnement des personnes dans le parcours de soin. Ce fût d'abord le fait de personnes atteintes par des pathologies graves voire incurables nécessitant un traitement lourd et à vie, comme les séropositifs ou les personnes atteintes de certaines hépatites.

²⁵³ Voir par exemple Guillaume Calafat, « Expertise et compétences: procédures, contextes et situations de légitimation », *Hypothèses*, 1 (14), 2011.

²⁵⁴ Voir, sur le processus d'expertise et les rapports de pouvoir, l'analyse proposée par Emmanuel Henry dans le secteur de la santé au travail, dans Emmanuel Henry, « Nouvelles dynamiques de savoirs et permanence des rapports de pouvoir. L'impact – limité – des transformations – importantes – de l'expertise en santé au travail », *Revue française de science politique*, vol. 61, n°4, 2011.

²⁵⁵ Guillaume Calafat, *art cit.*

expertise, le requérant lui attribue alors la nature d'expert²⁵⁶. Dans le cas qui nous intéresse ici, la sollicitation de PISA par les décideurs politiques français tend à valoriser les traits d'expertise qui sont propres aux agents de l'OCDE. Sur le principe les statuts professionnels des évaluateurs nationaux et internationaux sont proches du fait de la proximité des principes qui régissent le fonctionnariat national et international, et notamment le principe clé d'indépendance du pouvoir politique nécessaire à la recherche de l'intérêt collectif de l'action publique, nous y reviendrons. Les dimensions de l'expertise à mettre en avant sont alors à chercher à la fois dans le type de formation initiale suivi par les agents, et dans leur capacité à aborder la production de savoirs mobilisables pour l'action dans une perspective proche de celle du monde de la recherche, construisant des méthodes d'évaluation et des outils nouveaux remettant en cause les pratiques traditionnelles plus que s'inscrivant dans la continuité des modes de travail de son administration.

Dans le cas de nos évaluateurs, le crédit accordé aux résultats de PISA tient non seulement à l'usage politique que peuvent en faire les gouvernements qui y participent, mais aussi au crédit dont bénéficient les équipes qui y travaillent auprès de ceux qui en utilisent les productions. Plus que des compétences, ce sont des qualités d'experts qui sont allouées aux équipes de l'OCDE par nos interlocuteurs français, qu'ils soient au sein du gouvernement ou salariés de la DEPP. Cette même expertise est également revendiquée par les équipes de la DEPP en charge de suivre PISA parmi toutes les évaluations dont elles ont la charge. Les réunions de préparation de PISA sont alors des lieux et des moments où échangent des homologues, nationaux pour certains, internationaux pour les autres, experts dans leur domaine. Lieux de confrontation de savoirs, ils concrétisent dans un même espace-temps la controverse scientifique²⁵⁷ portant sur les sujets débattus. La difficulté de faire émerger notamment une définition unanime des résultats de l'éducation et des méthodes d'évaluation pouvant y être

²⁵⁶ Isabelle Backouche, "Introduction", numéro spécial "Devenir expert", *Genèses*, 70, 1, 2008, p.2.

²⁵⁷ Philippe Roqueplo, « L'expertise scientifique, consensus ou conflit ? », *La terre outragée, les experts sont formels*, Autrement, série sciences en société, n°1, 1992.

appliquées, laisse voir l'ampleur des débats que cela peut susciter, et la place prépondérante du conseil de l'expert pour rechercher les items les plus neutres et représentatifs possible du niveau des élèves dans les dizaines de pays participants. Le terme d'« expert » est d'ailleurs bien présent dans les écrits définissant PISA à sa création. Il désigne bien sûr les membres des « *groupes d'experts* » chargés de créer les items et les questionnaires, mais il est aussi employé pour décrire la qualité des membres du Conseil des Pays Participants. L'appendice du document donnant le rôle opérationnel du conseil des pays participants indique ainsi dès ses deux premiers points:

- « 1. Le Conseil des pays participants (CPP) aura pour tâche de suivre et d'évaluer le Programme de travail du Programme décentralisé chargé de mettre en œuvre la Stratégie pour l'élaboration périodique d'indicateurs sur les résultats des élèves, décrite dans le document DEELSA/ED/CERI/CD(97)4.
2. Il sera composé d'experts nationaux en matière d'évaluation des connaissances au nombre d'un expert par pays participant. »²⁵⁸

Du côté des équipes de l'OCDE, peu de qualificatifs apparaissent dans les documents officiels de PISA. En parcourant les volumes PISA et notamment les sections précisant les différents contributeurs, on peut toutefois lire qu'ils sont cités à la suite des « *experts et institutions travaillant au sein du Consortium PISA* »²⁵⁹ comme étant leurs collaborateurs dans le cadre de l'évaluation, aux côtés des pays participants (et donc des nombreux experts précités), et de l'UNESCO. S'ils ne se définissent pas officiellement dans les communications officielles comme des experts de l'action éducative, ce statut leur est cependant régulièrement accordé par ceux qui reçoivent leurs publications, et notamment les agents de la DEPP avec lesquels ils travaillent.

²⁵⁸ OCDE, *Résolution du conseil instituant un programme d'élaboration périodique d'indicateurs sur les acquis des élèves*, 15 décembre 1997, C(97)176/FINAL, p. 3.

²⁵⁹ OCDE-UNESCO-UIS, *Literacy Skills for the World of Tomorrow, further results from PISA 2000*, 2003, p. 4.

« Nous on est responsable de démultiplier PISA au niveau des pays. Les experts de l'OCDE, ils coordonnent. Ils sont garants de la publication des résultats. »²⁶⁰

« Entre les réunions du PISA Governing Board auxquelles je participe, il y a des réunions avec le secrétariat général, on échange quand même avec ceux qui mènent l'évaluation, ils ont l'expertise. »²⁶¹

L'expertise allouée à l'OCDE permet de garantir la validité à la fois du processus de construction de l'évaluation à laquelle les agents de la DEPP participent, et des indicateurs en eux-mêmes fournis aux décideurs publics des pays participants. Ce travail conjoint d'expert n'est pas exempt de rapports d'influence, visant à renforcer la légitimité des productions réalisées par chacun dans l'accompagnement de l'action éducative. Nous verrons que ces rapports d'influence se manifestent par la recherche du positionnement le plus en adéquation possible avec ce qu'ils perçoivent être les attentes de ceux amenés à utiliser les savoirs produits pour orienter les politiques éducatives.

a) La centrale indépendance de l'expert-fonctionnaire

Une fois constaté la revendication ou la reconnaissance du statut d'expert aux agents en charge de PISA à l'OCDE et aux agents de la DEPP, nous voulons observer la façon dont en pratique ces derniers conçoivent les dimensions clés de leur expertise, et les mettent plus ou moins en valeur afin de légitimer leurs productions. Le statut de fonctionnaire d'abord dont ils disposent tous est soumis à plusieurs règles théoriques qui donnent un fondement légal à l'expertise des évaluateurs de l'OCDE et de la DEPP. Les grands principes qui régissent le fonctionnariat sont en effet similaires quelle que soit l'envergure de l'organisation

²⁶⁰ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 21 mars 2019.

²⁶¹ Entretien avec Ronan Vourc'h, Adjoint au chef de bureau de l'évaluation des élèves, DEPP, 22 juin 2020.

dans laquelle ils sont impliqués. Parmi eux, l'indépendance de toute pression politique nationale ou internationale est centrale pour asseoir leur statut d'expert, et est citée dans les textes fondateurs des fonctionnariats aussi bien français que de l'OCDE.

Dans le cas des fonctionnaires français, leurs devoirs sont énoncés à l'article 25 de la loi Le Pors²⁶²: "*Le fonctionnaire exerce ses fonctions avec dignité, impartialité, intégrité et probité. Dans l'exercice de ses fonctions il est tenu à l'obligation de neutralité.*" Ces cinq principes se retrouvent à peu de choses près dans les textes qui énoncent les règles déontologiques applicables par les agents de l'OCDE. C'est une convention qui régit leur statut. Ils ne sont pas fonctionnaires nationaux détachés auprès d'une organisation internationale, et donc soumis aux textes de leur pays, mais ils se doivent de remplir un certain nombre de qualité proches de celles énoncées dans les textes de loi déterminant les conditions d'exercice de la fonction publique française.

Ils ont d'abord, et ce sera fondamental dans la façon dont s'organisent les relations entre les équipes de l'OCDE, et notamment de PISA, et les représentants des pays qui participent à l'évaluation, une indépendance statutaire²⁶³, qu'ils soient contractuels – avec des missions de plus ou moins courtes durées, allant de quelques mois à plusieurs années en cas de renouvellement des contrats - ou membres permanents de l'organisation. Le statut de fonctionnaire international est en cela particulier qu'il ne peut, et c'est le cas lorsque l'on parle de l'OCDE notamment, se voir appliquer de façon automatique les règles de la fonction publique. L'accord des pays participants aux instances internationales doit être recueilli pour que soient appliquées les règles de droit public, y compris lorsque les missions des organisations internationales se rapprochent de missions d'intérêt général du secteur public²⁶⁴. D'autre part, une fois acquis ce statut de fonctionnaire, les règles propres à la

²⁶² Loi n°83-634 du 13 juillet 1983 portant droits et obligations des fonctionnaires.

²⁶³ Finck François, « L'indépendance statutaire des agents du conseil de l'Europe », *Revue française d'administration publique*, 2, n° 126, 2008.

²⁶⁴ Alain Plantey, François Lorient, Chapitre 1. Nature et champ de la fonction publique internationale, *Fonction publique internationale: Organisations mondiale et européennes*. Paris; CNRS Editions, 2005.

nature internationale des missions doivent être constituées et ne peuvent être celles du pays hôte. L'organisation internationale ne peut être placée en situation de dépendance vis-à-vis du pays qui en héberge les bureaux, et progressivement ont été édictées des règles de fonctionnement propres, à mesure que le nombre des organisations internationales augmentait²⁶⁵. Parmi ces règles, l'indépendance des agents est citée en bonne place. La convention de l'OCDE énonce ainsi à son article 2 :

« Les agents ne sollicitent ni n'acceptent de directives d'aucun pays Membre de l'Organisation ni d'aucune source extérieure à l'Organisation. Sauf s'ils y ont été autorisés, ils ne sollicitent ni n'acceptent aucune :

- i) gratification ou avantage en raison de l'exercice de leurs fonctions ou de leur statut d'agent de l'Organisation ;
- ii) distinction honorifique ; ou
- iii) rémunération »²⁶⁶

L'indépendance entre l'expert et le décideur est une dimension fondamentale de légitimité de l'expert²⁶⁷. L'objectif d'intérêt public des prises de décision implique que le conseil donné ne soit pas pré-orienté par le but souhaité par le décideur. En d'autres termes, on ne cherche pas avec l'expertise à montrer ce que l'on souhaite faire, mais on cherche à recueillir tous les éléments permettant de prendre une décision la plus impartiale possible. De la même façon, les experts sollicités ne doivent trouver intérêt ni dans un lien entretenu avec les commanditaires de l'expertise, ni dans le résultat que la décision politique suivant leur expertise pourrait avoir²⁶⁸.

Dans le cas de nos agents de l'OCDE, l'indépendance du gouvernement du pays dont ils sont originaires permet non seulement au niveau individuel de pouvoir « faire son travail » sans risquer de subir des pressions politiques lorsque les sujets qu'ils traitent présentent un intérêt majeur pour leur pays d'origine. Il

²⁶⁵ Alain Plantey, *Principes et développement du processus international de décision, L'avenir des organisations internationales*, Paris, 1983.

²⁶⁶ OCDE, *Statut, règlement et instructions applicables aux agents de l'organisation*, actualisation de mars 2020, article 2 p.8.

²⁶⁷ Renaud Bouvet, *art. cit.* p. 64.

²⁶⁸ Olivier Leclerc, *L'indépendance de l'expert*. K. Favro (coord.). *L'expertise : enjeux et pratiques*, Lavoisier/Tec&Doc, 2009.

permet aussi à l'OCDE dans son ensemble de garantir à ses lecteurs une qualité de production par la neutralité statutaire de leurs travaux. Les pressions gouvernementales qui peuvent demeurer, car les organisations nationales sont bien des terrains d'influence on le verra, sont alors -diplomatiques entre des entités politiques, dans le cadre par exemple de réunions plurilatérales faisant l'objet de comptes-rendus. Les pressions financières possibles sont réduites au minimum par un système de contribution obligatoire selon le PIB des membres, et de participation volontaire en cas de demande d'intégrer tel ou tel programme ou telle ou telle évaluation. Dans la pratique, les agents de la Direction de l'Education de l'OCDE que nous avons pu rencontrer font tous part de l'absence de pression directe auprès d'eux exercée par les équipes de leur pays d'origine travaillant sur PISA. Ils sont avant tout agents de l'OCDE avant d'être de telle ou telle nationalité et ne sont, sauf rares exceptions, jamais détachés de la fonction publique nationale vers l'organisation internationale, comme cela peut par exemple être le cas dans certaines institutions européennes.

Autres principes d'action écrits noir sur blanc dans la convention de l'OCDE, qui se rapprochent des principes énoncés par la loi Le Pors : l'intégrité, l'objectivité et l'impartialité.

« a) Les agents :

ARTICLE 3

- i) remplissent leurs fonctions avec la plus haute intégrité et loyauté ;
- ii) se comportent avec objectivité et impartialité et évitent tout conflit d'intérêts, ou apparence de conflit d'intérêts, dans l'accomplissement de leurs fonctions ;
- iii) gèrent avec soin les ressources de l'Organisation dont ils sont responsables ;
- iv) n'utilisent pas les ressources de l'Organisation à leur profit personnel ou à celui de tiers ;
- v) divulguent toute fraude, corruption et mauvaise utilisation des ressources de l'Organisation ;
- vi) ne font pas valoir la position qu'ils occupent au sein de l'Organisation, ses nom, logo ou tout renseignement obtenu dans l'exercice de leurs fonctions pour

obtenir un avantage indu à leur profit ou à celui de tiers, ou pour toute autre fin inappropriée.

b) Les agents ne doivent subir aucune forme de représailles ou aucun préjudice en raison de la divulgation de faits conformément aux dispositions de l'article 3

a) v) du présent statut et de toute instruction connexe. »²⁶⁹

Dans le cas des évaluations qui nous intéressent, fonctionnant par tests standardisés, l'objectivité des résultats est censée être atteinte autant par l'objectivité de ceux qui les conçoivent et qui créent les indicateurs finaux, que grâce à l'utilisation de méthodes statistiques particulièrement robustes accompagnées d'une construction par tous les participants des items passés aux élèves, on l'a vu.

Néanmoins, au-delà de la théorie des textes qui accordent au fonctionnariat des principes d'action proches de ceux que l'on attend d'un expert, l'indépendance des deux administrations est régulièrement mise en cause. Les principes d'objectivité et d'impartialité, pourtant inscrits en tête de sa convention, sont souvent remis en cause lorsqu'il s'agit des travaux de l'OCDE, et donc par dérivation de ses directions et de la direction de l'éducation. L'OCDE est en effet régulièrement accusée d'obéir aux volontés des plus gros contributeurs financiers et des initiateurs historiques de l'organisation, les Etats-Unis. Dans le cas de PISA, c'est la forme prise par les évaluations qui a suscité de la méfiance au début de la vie du Programme, et en particulier du côté des lecteurs français. Le test par QCM, très utilisé aux Etats-Unis pour évaluer les élèves, a notamment été considéré dès le premier cycle comme étant peu adapté au système français, et ne permettant pas de rendre compte de façon objective des résultats des élèves de 15 ans²⁷⁰.

²⁶⁹ OCDE, *Statut, règlement et instructions applicables aux agents de l'organisation*, actualisation de mars 2020, article 3, p.10.

²⁷⁰ Voir chapitre 4 pour plus de précisions.

Dans le cas de la DEPP, c'est l'intégration dans le ministère de l'éducation et la nomination en conseil des ministres de son directeur ou de sa directrice qui donnent à leurs interventions auprès des équipes PISA-OCDE une tonalité potentiellement proche du pouvoir politique. La distinction entre le politique, donnant les orientations générales de politique éducative, souvent proche de la direction de la DEPP et influant sur le budget de la direction, et l'administration orientée vers la recherche, libre de son programme annuel de travail, a pu apparaître comme peu claire aux yeux de certains décideurs. Cela a pu donner lieu, comme sous le Ministère Allègre, à un délitement du lien entre le ministre et sa direction²⁷¹. Cela a pu également, à l'opposé, donner lieu à des accusations de manipulation, par un certain nombre de syndicats d'administrations en 2011, des données de la DEPP pour valoriser l'action ministérielle²⁷². L'attribution à la DEPP du statut de service statistique ministériel depuis 2016, respectant les principes édictés au niveau européen par le Parlement et le Conseil²⁷³ concernant la statistique publique a entériné l'ensemble de ses devoirs, au premier rang desquels l'indépendance et l'impartialité des statistiques.

Ce doute existant en toile de fond sur l'absence d'indépendance des évaluateurs a été notamment repris à dessin par les détracteurs des travaux d'évaluation et en particulier de PISA. L'enjeu dans le cadre de PISA a été alors de réaffirmer la valeur de l'expertise produite non seulement en rassurant sur l'indépendance des statistiques produites et leur objectivité, par le biais d'une méthodologie reconnue par sa robustesse et aux multiples contributeurs de tous pays, mais aussi par un recrutement censé garantir les compétences des membres

²⁷¹ Voir chapitre 1 pour plus de détails.

²⁷² Tract commun CGT, CFDT et SNUTETE FSU de l'administration centrale du Ministère du travail, de l'emploi et de la santé (DARES et DREES), CGT Educ'action centrale du Ministère de l'Education nationale et Ministère de l'enseignement supérieur et recherche, SYGMA FSU du Ministère de l'Agriculture et de la pêche (SCEES), CGT Equipement-environnement de l'Administration Centrale du Ministère de l'Ecologie, du Développement durable, des Transports et du Logement, CGT, CFDT, SUD de l'Insee, CGT de Pôle Emploi, « L'indépendance de la DEPP (Direction de l'Evaluation, de la Prospective et de la Performance) mise à mal : le débat faussé sur l'Education nationale, publié le 2 novembre 2011.

²⁷³ Règlement CE n°223/2009 du Parlement européen et du Conseil européen du 11 mars 2009 relatif aux statistiques européennes, modifié par le règlement UE 2015/759 du Parlement européen et du Conseil européen du 29 avril 2015.

des équipes qui y travaillent. Si le principe du recrutement du candidat le plus compétent va de soi, c'est ici ce que les organisations valorisent comme compétences qui nous intéressent pour ce qu'il témoigne de ce qu'elles entendent mettre en avant au moment où est évaluée l'expertise de leurs équipes.

Au-delà des principes qui régissent leur travail et qui déterminent donc en partie la qualité du produit qu'ils fournissent aux décideurs publics, le recrutement des membres des équipes de PISA-OCDE et de la DEPP se fait au départ selon les mêmes grands principes caractéristiques de la fonction publique. Dans les deux organisations sont mis en place des processus de recrutement censés garantir du haut niveau de compétence des personnes intégrant les équipes, haut niveau établi selon des critères propres à chacune qui se rejoignent partiellement.

Dans le cas de la DEPP, l'équipe en charge de l'évaluation PISA a été recrutée selon deux logiques dont l'importance de l'une par rapport à l'autre a évolué au fur et à mesure des cycles. Le cœur de l'équipe, et en particulier les premiers National Project Managers (NPM) ou membres du Conseil des Pays Participants, a été recruté selon les principes classiques de recrutement dans la fonction publique. Ils ont d'abord pour la plupart passé avec succès un concours de la fonction publique qu'il s'agisse d'un concours pour accéder à un poste d'enseignement, d'un concours administratif pour intégrer la fonction publique d'Etat ou d'un concours permettant de faire ses premières expériences au sein d'instituts nationaux réputés très exigeants comme l'INSEE. Les recrutements ont été menés ensuite selon les processus classiques d'appel à candidature, suivis d'entretiens d'embauche avec à la fois un membre de la direction et un représentant des ressources humaines.

L'évolution des profils est intéressante à noter, de plus en plus spécialisés sur la méthode d'évaluation employée notamment par PISA. Dans le cas de la DEPP, les compétences recherchées sont plurielles et les équipes constituées pluridisciplinaires. La logique interne de constitution des équipes ne varie pas selon les projets sur lesquels les équipes travaillent. La constitution de l'équipe

PISA a suivi les mêmes règles que toutes les autres, à savoir l'importance de combiner à la fois une compétence technique, celle de la maîtrise avancée des méthodes de l'évaluation et/ou des outils de traitement statistique, et celle de la « compétence métier », généralement acquise au cours d'années d'enseignement ou d'une formation en lien avec les pratiques pédagogiques. D'années en années et au fil de l'apparition dans les lieux de formation de modules consacrés aux méthodes d'évaluation, les profils se sont toutefois spécialisés.

Lors de nos entretiens, au-delà de la longue expérience professionnelle des interlocuteurs rencontrés, tous ont tenu à indiquer la thématique de leur formation initiale, quelle que soit son ancienneté. Ont été souvent mises en avant les formations en statistique ou très spécifiquement sur l'évaluation des politiques publiques - voire même des politiques éducatives. La « connaissance du terrain »²⁷⁴ a également été valorisée par plusieurs NPM qui ont eu des expériences d'enseignement en parallèle d'un cursus leur conférant des compétences techniques pour comprendre le travail des indicateurs et pour pouvoir échanger « d'égal à égal » avec leurs interlocuteurs de l'OCDE lors des réunions de préparation des différents cycles de PISA. La combinaison des deux dimensions, enseignement et maîtrise de la technique d'évaluation, apparaît souvent dans la description qu'ont pu faire les NPM ou les responsables d'équipe de la DEPP de leurs parcours.

« Alors j'ai été recruté comme instituteur à l'époque, j'ai passé un bac en mathématiques, après j'ai fait l'école normale d'instituteur. J'ai enseigné à l'étranger avec des adultes, et au retour j'ai repris des études. J'ai fait une maîtrise en sciences de l'éducation, un doctorat en linguistique. Au cours de ces études j'ai eu comme professeur Claude Thélot, il m'a proposé de faire un mémoire sur des problématiques que rencontrait la DEPP pour prendre en charge les élèves non francophones. J'ai aussi fait une licence en statistique après ça, il m'a proposé de travailler au sein de la DEPP. (...) Dans l'éducation c'est très important d'avoir les deux, parcours universitaire et terrain. On peut arriver par le terrain ou par l'autre, dans

²⁷⁴ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 21 mars 2019.

les deux sens, mais avoir les deux c'est très important pour saisir les enjeux. »²⁷⁵

« Aujourd'hui je suis adjoint au chef de bureau de l'évaluation des élèves, j'y suis depuis un peu plus d'un an, je m'occupe de l'orientation macro de PISA. Mon parcours, je suis ici depuis 2011, j'ai fait un écart par l'enseignement supérieur pendant deux ans, et je suis revenu en tant qu'adjoint. J'ai une formation en sociologie quantitative, dans le champ de l'éducation. Je gère aujourd'hui du coup les traitements de données, je représente la DEPP au Pisa Governing Board. »²⁷⁶

« Je suis chargé d'étude au bureau de l'évaluation des élèves. Je suis en demi-poste, c'est-à-dire que je continue d'enseigner. Je suis chargé d'étude ici sur les dossiers internationaux en mathématiques, et je travail en réseau d'éducation prioritaire la moitié de mon temps. J'ai été recruté en 2009 pour prendre en charge le groupe de conception DEPP pour les items en maths pour le cycle 2012, je connais bien la réalité de l'enseignement du coup c'était intéressant. Je représentais la France, j'étais co-National Project Manager dans les meetings internationaux, puisqu'on avait l'habitude d'envoyer comme représentants des spécialistes du domaine majeur. »²⁷⁷

Le fait de disposer d'une expérience d'enseignement est fortement valorisé par certains membres de l'équipe car elle suppose une meilleure connaissance des conditions d'apprentissage des élèves et permet de compléter une formation théorique en statistique ou en sciences sociales. Des enseignants sont d'ailleurs recrutés en demi-postes spécifiquement pour apporter leur regard sur la conception d'items dans leur discipline et piloter les groupes de travail nationaux constitués de personnes du corps enseignant et des corps d'inspection. S'ils sont recrutés lorsque leur discipline se trouve être la majeure de l'évaluation PISA en cours lors de leur recrutement, ils continuent souvent à travailler à temps partiel sur les évaluations suivantes. Leur rôle est clé au sein de l'équipe, au point qu'ils

²⁷⁵ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 18 mars 2019.

²⁷⁶ Entretien avec Ronan Vourc'h, Adjoint au chef de bureau de l'évaluation des élèves, DEPP, 22 juin 2020.

²⁷⁷ Entretien avec Franck Salles, Chargé d'étude sur l'évaluation des élèves, National Project Manager, DEPP, 11 juillet 2017.

sont considérés pour certains comme des « co-NPM », présents lors des réunions internationales avec un travail de représentation qu'ils revendiquent pleinement.

L'expertise des évaluateurs de la DEPP est au départ surtout accordée par leur expérience professionnelle, que ce soit par la connaissance actuelle et « par le vécu » des enjeux de l'évaluation des compétences des élèves ou par le gage de qualité que confère le statut de membre de la fonction publique acquis par concours national, ou par leur maîtrise de la statistique appliquée. Au départ, rares sont ceux qui peuvent se définir comme des experts en évaluations psychométriques spécifiquement appliquées au champ éducatif²⁷⁸. Au fur et à mesure des cycles, la nécessité de monter en compétences en psychométrie, et de pouvoir la sanctionner par un diplôme du plus haut niveau, apparaît. Certains agents de la DEPP, pourtant impliqués au premier rang dans l'évaluation PISA pour la France dès le départ, complètent ainsi leur cursus en formation continue en obtenant un doctorat. C'est le cas très particulièrement de Thierry Rocher, actuel chef du bureau pour l'évaluation des élèves et entré à la DEPP à la sortie de ses études à l'École Nationale de la Statistique et de l'Analyse de l'Information à la fin des années 1990. Il est chargé dès les tous premiers cycles de suivre PISA et en sera le premier NPM pour la France. Après avoir travaillé sur PISA pendant plusieurs cycles, il soutient une thèse en psychologie et psychométrie à l'Université de Nanterre soutenue en 2013 où il étudie les méthodes de mesure des compétences en compréhension de l'écrit, et notamment les relations entre les modèles psychométriques et la psychologie cognitive. Ce diplôme de doctorat, très reconnu certes au niveau national même si peu demandé dans la fonction publique à l'époque, est en revanche très prisé des agents de l'OCDE qui en font une marque fondamentale de la démarche intellectuelle qu'ils prônent, plus proche de la recherche que de l'application de modèles statistiques à des questionnements en politique éducative.

²⁷⁸ En particulier, Bruno Trosseille, en charge du bureau des évaluations des élèves au début des années 2000 et qui a été NPM lors des premiers cycles de l'évaluation avec Thierry Rocher, a une formation de conseiller d'orientation psychologue et est très au fait des principes des tests psychométriques qu'il a abordé lors de sa formation et qui ont suscité son intérêt personnel.

Dans le cas de l'OCDE, un processus de recrutement tout à fait exigeant et en plusieurs étapes est appliqué. Il ne se déroule pas comme dans d'autres organisations internationales qui sont composées de fonctionnaires nationaux qui suivent un processus de recrutement sur concours ou qui peuvent même être détachés de leur administration nationale. Pour ce qui est de l'équipe PISA, les membres ont été recrutés pour leurs compétences théoriques jugées de « haut niveau » et pour leur expérience des évaluations internationales, qu'elle ait été acquise en interne avec un système de « gravisement des échelons » au fur et à mesure des éditions de PISA, ou dans d'autres organisations internationales pratiquant l'évaluation ou l'analyse statistique. En termes de diplôme souhaité, il est d'ailleurs indiqué lors des recrutements que la formation demandée est bien celle d'un très haut niveau universitaire. Finalement, la valeur de diplômes des grandes écoles « à la française » n'est pas spécifiquement reconnue, au-delà du niveau scolaire qu'elle procure. L'OCDE se place en ce sens à un niveau international en dehors du système de « grande école » qui n'a que peu de sens dans une organisation où des dizaines de pays sont présents à travers les agents qui y travaillent. En revanche, les palmarès universitaires sont considérés comme représentatifs de la qualité de maîtrise technique et de la force de travail des agents recrutés.

Le poids des diplômes, des parcours professionnels, et des compétences personnelles de chaque membre de l'équipe est mis en avant, y compris dans leur description sur le site de l'organisation. Un « *Who's who* » est mis en ligne sur le site de PISA²⁷⁹ reprenant pour chaque personne y travaillant une courte biographie, à la manière des dictionnaires biographiques qui existent depuis des décennies dans plusieurs pays occidentaux, et qui donnent les biographies des « personnes qui comptent », « des meilleurs talents »²⁸⁰ pour ne citer que l'édition française. Chaque membre de PISA-OCDE est ainsi décrit par son poste, son ancienneté dans l'organisation, son parcours au sein des plus grandes institutions

²⁷⁹ <https://www.oecd.org/pisa/contacts/whoswhoinpisa.htm>

²⁸⁰ Voir la page « Qui sommes nous » du site du Who's who français, www.whoswho.fr, qui présente le dictionnaire comme un recueil sélectif des personnalités françaises « remarquables ».

de son pays d'origine ou internationales, et par son (ou le plus souvent ses) diplômes des universités les plus réputées du monde. Le tout sans oublier les nombreuses langues parlées mentionnées en fin de présentation. Le classement des biographies est fait par ordre alphabétique et non par un quelconque ordre hiérarchique en termes de poste ou encore d'ancienneté. Les derniers arrivés se trouvent ainsi entre les chefs de division et les chefs de projets seniors. Cette présentation horizontale des personnes est par ailleurs presque décontractée à plusieurs égards. La présentation est faite par le prénom, y compris lorsqu'il s'agit des responsables de publication. Une grande importance est accordée à la maîtrise des langues comme signe de curiosité, d'ouverture sur le monde et d'excellence intellectuelle. Les plus hauts placés sur l'échelle hiérarchique de l'équipe PISA valorisent ainsi leur motivation et leur investissement à continuer à apprendre en terminant leur biographie en précisant qu'ils étudient une langue étrangère, même si leur CV est reconnu par leurs pairs.

L'équipe est constituée d'agents recrutés dans l'organisation au terme d'un parcours en plusieurs étapes. Certains sont entrés en stage après leur diplôme universitaire et ont eu plusieurs contrats de consultant interne – des contrats à durée déterminés souvent d'une durée de quelques mois et renouvelables – avant de passer les concours pour devenir agent permanent de l'OCDE et obtenir le statut de fonctionnaire international. D'autres ont été recrutés directement après de premières expériences professionnelles dans des organisations internationales travaillant sur des sujets proches, en passant les concours dès le départ. Ils fonctionnent sur un principe d'écrit puis d'oral, avec un système d'entonnoir aboutissant à une liste de cinq personnes pouvant obtenir un poste permanent, appelée « shortlist ». Une première étape de tests statistiques et d'interprétation des données est organisée, suivie pour ceux qui sont retenus d'un passage devant un panel de plusieurs personnes représentant les ressources humaines de l'organisation, les autres directions, la direction de l'éducation et l'équipe concernée par l'ouverture du poste. Les cinq derniers sont donc mis sur une liste leur permettant de candidater lors d'une prochaine ouverture de poste, sans devoir repasser par tout le processus. Les équipes PISA-OCDE ont ainsi d'abord

été resserrées à une dizaine de personnes dans les premiers cycles, toutes n'intervenant pas dans l'intégralité de la publication mais pouvant apporter des compétences spécifiques sur certains indicateurs relevant de leur maîtrise. Au fur et à mesure des cycles l'équipe a été étoffée de contributeurs nombreux apportant une « *assistance analytique (...) supplémentaire* »²⁸¹, dont le nombre est à la mesure de l'enrichissement de la publication passée en quelques années de un à six volumes.

Les recrutements au sein des deux équipes se font finalement davantage en fonction des profils jugés intéressants et de haut niveau des personnes qui postulent à un poste qu'en fonction du statut des personnes qui postulent. Si dans les deux cas un noyau d'agents ayant le statut de fonctionnaire dans le cas de la DEPP ou de membre permanent de l'organisation dans le cas de l'OCDE constitue l'équipe en charge de l'évaluation internationale, cette dernière est très régulièrement complétée par des personnes permettant d'apporter des compétences complémentaires nécessaires pour travailler sur des thématiques précises des différents cycles. Cette ouverture vers l'extérieur n'obéit pas totalement à la même logique dans les deux organisations. Dans le cas de l'OCDE, les recrutements d'agents venus de l'extérieur, y compris pour travailler sur l'ensemble de la publication et pas uniquement sur certaines thématiques pouvant nécessiter des compétences spécifiques, obéit à une logique de pluralité des points-de-vue. La diversité des profils est mise en avant afin de travailler en combinant des visions du monde plurielles, de créer des environnements de travail plus stimulants, d'apporter des savoirs et des compétences différents, et donc de garantir la pertinence des travaux de l'OCDE. Pour la DEPP, mis à part les enseignants sollicités ponctuellement pour accompagner la préparation des évaluations correspondant à leur discipline, il s'agit davantage de pourvoir des postes qui sont complexes à pourvoir en interne, avec la perspective éventuellement pour les contractuels de passer les concours de la fonction publique.

²⁸¹ OCDE (2015), *Résultats de PISA 2012 : Les élèves et l'argent (Volume VI) : les compétences en culture financière au XXIe siècle*, PISA, Editions OCDE, p.7.

b) L'enjeu de l' « expertise créative »

La sollicitation par l'OCDE d'experts ponctuels, qu'ils soient internes ou externes, dans la phase d'analyse des résultats des tests PISA et de construction des indicateurs qui seront mis en avant dans les publications est une façon d'asseoir la dimension de recherche de ses travaux. Cette volonté d'innover, de créer de nouveaux indicateurs faisant référence au sein de la discipline statistique, est mise en avant par l'organisation internationale comme une vraie marque d'expertise. Le module Stata Repest par exemple, conçu par deux agents de l'OCDE titulaires de doctorats de l'EHESS, Francesco Avvisati et François Keslair, utilisé pour PISA, a ainsi donné lieu à des articles publiés par des universités américaines renommées²⁸² et dans des ouvrages spécialisés sous la direction d'universitaires travaillant sur l'évaluation des résultats de l'éducation. Cette mise en avant de la dimension de recherche des indicateurs de PISA construits par l'équipe de l'OCDE est centrale pour mieux saisir les rapports d'influence entre nos évaluateurs nationaux et internationaux. Ce que l'on pourrait appeler "l'expertise créative", où l'apport de connaissance pour l'action passe par la recherche permanente de l'indicateur le plus juste possible et par la mise en lumière de problématiques d'action publique objectivées par la statistique, est présentée par les pilotes de PISA à l'OCDE comme une condition sine qua non de la mission qui leur est confiée. Face à cette définition particulière de l'expertise mise en avant par l'OCDE, les évaluateurs de la DEPP peinent à s'inscrire dans la même démarche malgré le souhait exprimé de pouvoir davantage exploiter les résultats des élèves français à PISA pour aller plus loin dans leurs analyses.

L'équipe PISA-OCDE est dans une posture d'innovation technique notamment pour répondre au mieux aux multiples demandes des pays

²⁸² Francesco Avvisati et François Keslair, *REPEST, Stata module to run estimations with weighted replicate samples and plausible values*, Statistical Software Components, Boston College Department of Economics, 2020; François Keslair, *Analysing PIAAC Data with Stata*, dans Debora Maehler, Beatrice Rammstedt (dir.), *Large-Scale Cognitive Assessment. Methodology of educational measurement and assessment*, Springer Cham, 28 juillet 2020.

participants et à l'évolution des contextes de vie et des exigences de formation initiale d'un cycle à l'autre. Le principe de fonder les évaluations sur la capacité à répondre aux situations de la vie courante, plus que sur un stock de connaissances sous la forme d'un programme peu dynamique à court terme, implique cette plasticité des indicateurs rendant compte des résultats des élèves.

Cette posture de recherche et d'innovation permanente permet également à l'équipe PISA-OCDE se répondre aux critiques portées à l'encontre des organisations dédiées à l'évaluation des politiques publiques. Certains travaux pointent en effet le risque de voir s'institutionnaliser un corps d'experts sollicités de manière systématique, à titre d'activité principale, qui soit déconnecté de l'évolution de l'état des connaissances dans leur domaine, déconnecté du monde de la recherche, et donc moins à même de fournir des compétences actualisées dans un environnement de connaissances sans cesse mouvant²⁸³. Les équipes de PISA OCDE se dégagent de ce risque en mettant en avant leurs parcours de formation initiale universitaire de très haut niveau, souvent conclu par l'obtention d'un doctorat et/ou d'un cursus pluridisciplinaire. Le fait qu'ils soient dédiés à temps plein à la préparation des prochains cycles, la conduite du cycle en cours et l'analyse des résultats mis en forme par le consortium leur permet également de mener l'analyse le plus loin possible avec le temps qu'ils ont à leur disposition.

Au sein de la DEPP, l'organisation interne de l'équipe travaillant sur PISA ne lui permet pas de s'engager aussi pleinement qu'ils le souhaiteraient dans une démarche analytique plus poussée que les notes d'information qu'ils publient après chaque cycle. L'importance croissante de PISA dans les évaluations qu'elle se doit de mener pour remplir son rôle de service de statistique ministériel, se fait sans augmentation proportionnelle ni de ses effectifs ni de son budget. En interne, plusieurs de nos interlocuteurs ont fait part de leur amertume en voyant parfois les budgets importants alloués à PISA au détriment de leurs projets à l'échelle nationale, alors même que les marges de manœuvre dont dispose la DEPP dans

²⁸³ Renaud Bouvet, *art. cit.* p. 64.

la façon dont sont construits les cycles PISA lui donnent un rôle d'initiative assez limité. Ces marges se bornant souvent, dans les réunions, à se positionner pour ou contre une proposition faite par l'OCDE ou les autres pays participants sans pour autant pouvoir proposer une alternative qui lui paraîtrait plus intéressante. Si les NPM ont pu proposer des thématiques de travail ou reformuler des items qui leur paraissaient pouvoir induire un biais, peu de ces propositions ont finalement été retenues.

Malgré la spécialisation croissante des équipes de la DEPP travaillant sur PISA et le caractère de recherche revendiqué par certains membres lorsqu'ils parlent de leurs travaux, cette identité de « fonctionnaire chercheur » peine à être reconnue unanimement, que ce soit par les autres instances composant l'administration du ministère de l'Éducation nationale et par les ministres eux-mêmes. Bien que la DEPP ait l'autonomie théorique de son programme de travail, bien que dans le cadre de PISA elle soit laissée officiellement maîtresse de ses interventions auprès de l'OCDE pour orienter les choix faits dans la construction des tests et l'exploitation des résultats, bien qu'elle fournisse des analyses problématisées et pas seulement descriptive des données dont elle dispose par le biais de ses évaluations et enquête, bien qu'elle se joigne à des équipes de chercheurs pour mener certains travaux exploratoires, le rôle qui leur est attribué par leur administration et leur ministre n'a pas de dimension de recherche. Cela ne signifie pas pour autant qu'en pratique les équipes de la DEPP n'ont pas développé des liens avec la communauté scientifique par le biais d'appels d'offres par exemple visant à mener des travaux sur des thématiques de travail²⁸⁴ proches de la recherche – « l'investissement éducatif et son efficacité », ou encore « les disparités territoriales et régionales en éducation et formation » par exemple. La direction a également contribué au développement de sujets de recherche, d'ailleurs assez proches des réflexions menées dans le cadre de PISA et de l'enquête TALIS portant sur les conditions d'enseignement et menée

²⁸⁴ Jean-Claude Emin, « Réflexions sur le rôle de la Direction de l'évaluation et de la prospective (DEP) en France », *Éducation et sociétés*, vol 2, n°30, 2012.

également par l'OCDE, sur les facteurs d'efficacité des établissements scolaires²⁸⁵.

Cela ne semble pas avoir été repris par les acteurs au sommet du ministère de l'Education nationale et la description des activités de la DEPP paraît davantage relever de l'application de directives que d'une véritable démarche de recherche-action. Cela transparaît notamment dans les documents budgétaires de ces dernières années²⁸⁶ :

« La DEPP contribue à l'évaluation du système éducatif dans toutes ses dimensions : évaluation des élèves, des établissements, des territoires, de dispositifs ainsi que du système éducatif dans son ensemble. Elle mobilise à cette fin le système d'information statistique sur l'éducation, dont elle a la charge, et conduit une quinzaine d'opérations d'évaluation d'élèves chaque année. Aux traditionnelles évaluations sur échantillons, internationales ou nationales, sont venues s'ajouter, depuis les rentrées 2017 et 2018, des évaluations exhaustives conduites auprès de plusieurs générations d'élèves (plus de 800 000 élèves concernés pour chacune de ces opérations). L'objectif est de fournir aux professeurs des repères des acquis de leurs élèves, de doter les « pilotes de proximité » d'indicateurs leur permettant de mieux connaître les résultats des élèves et des unités de formation pour adapter leur action éducative ; enfin de disposer du profil des élèves à l'entrée de l'école, du collège ou du lycée pour construire, au niveau national, de nouveaux indicateurs (valeur ajoutée par exemple). »

La DEPP voit dans cette description pour l'année à venir ses missions bien cadrées, au nombre d'évaluations près, avec une base de travail « traditionnelle » dont elle doit s'acquitter. Concernant les données qu'elle sera amenée à utiliser pour ses travaux, elle « mobilisera » celles à sa disposition, et les éventuelles nouveautés comme l'évaluation de la valeur ajoutée des établissements scolaires sont déjà suggérées dans le descriptif de ses missions.

La DEPP serait à la fois maîtresse de son programme de travail tout en devant remplir une liste d'obligation prévue à l'avance et à laquelle

²⁸⁵ Jean-Claude Emin, *art. cit.*

²⁸⁶ Annexe budgétaire au Projet de loi de finance 2019 Enseignement scolaire, Programme 214, Soutien de la politique de l'éducation nationale, Action n°02 Evaluation et contrôle.

correspondrait son budget pour l'année. Elle serait par ailleurs laissée autonome dans sa gestion de PISA et son intervention auprès de l'OCDE lors de la mise en œuvre des cycles et la réception des résultats, mais aurait en pratique peu d'influence sur les choix effectués. Tout cela en ayant peu de temps et de moyens pour exploiter les résultats fournis par l'OCDE pour la France. Finalement on voit là toute la difficulté de la construction de la légitimité de l'expertise des évaluateurs de la DEPP mis face à leurs homologues internationaux. Leur positionnement entre l'expert indépendant et le représentant national est complexe. Les NPM sont envoyés comme experts français dans les échanges avec l'OCDE sans être outillés par le politique pour porter leur parole et les orientations qu'ils ont déterminées pour les politiques éducatives à venir. Ils sont dans le même temps chargés de porter avec eux leurs compétences et leur analyse des sujets abordés, travaillés avec leur équipe pluridisciplinaire au préalable, et ils sont chargés de « contribuer à l'élaboration d'un « jugement de vérité » » sur les différents aspects des problèmes abordés²⁸⁷. On voit là apparaître un hiatus entre la façon dont les travaux de la direction sont décrits dans les documents de pilotage de l'activité du ministère et dont le ministère souhaite qu'ils interviennent auprès de l'OCDE, et la mise en place de ces principes d'action en pratique dans le cadre de la participation de la France à PISA. Alors que la DEPP peine à faire entendre ses propositions auprès de ses partenaires internationaux travaillant à PISA, l'équipe de PISA-OCDE se pose dans le même temps en « entrepreneur de problème public²⁸⁸ » pour placer au centre des débats les problématiques que l'OCDE considère les plus importantes à travailler en prévision des évolutions sociales et économiques à venir.

²⁸⁷ Roqueplo, *art cit.* p. 13.

²⁸⁸ La notion est dérivée des travaux d'Erik Neveu sur l'entrepreneur de problèmes, qu'il développe à partir du concept d'entrepreneur de cause développé par Howard Becker et qu'il analyse notamment dans son ouvrage *Sociologie politique des problèmes publics*, Armand Colin, 2015, et en particulier dans le chapitre 2 intitulé "Identifier".

c) Le pouvoir d'entrepreneur de problèmes publics au cœur de la constitution des rôles de nos évaluateurs

Par leur positionnement de recherche, voire de conseil aux nouveaux gouvernements en place, l'équipe PISA-OCDE se pose en entrepreneurs de problèmes publics, au sens où elle met en lumière certaines situations qu'elle contribue à faire percevoir comme problématiques. Les indicateurs qu'elle produit ne sont pas neutres. Lorsque l'OCDE choisit de mettre l'accent sur telle ou telle problématique en proposant aux pays de les évaluer lors des prochains cycles de PISA, et travaille à la construction d'indicateurs ad hoc, elle incite les pays à s'interroger sur la façon dont ils les traitent au niveau national. En reliant par exemple la performance aux tests PISA et d'autres caractéristiques aussi bien individuelles des élèves (l'origine socio-économique des élèves par exemple) qu'environnementales (liées aux établissements fréquentés et bien sûr, sans que les facteurs de causalités ne soient clairement identifiés, les pays dans lesquels les élèves suivent leur scolarité), les indicateurs produits amènent les responsables politiques, les chercheurs, l'opinion publique, à prendre connaissance des marges d'amélioration qui sont envisageables en travaillant les problématiques qui émergent. C'est d'ailleurs bien l'objectif énoncé par Andreas Schleicher lorsqu'il est interrogé sur les nouveaux indicateurs qui sont créés par son équipe pour PISA, pour pouvoir donner à voir les besoins liés aux compétences indispensables pour être un adulte dans la société de demain. PISA se doit non seulement de montrer les problématiques actuelles parfois invisibles aux yeux des décideurs publics, ou volontairement mis de côté selon les tendances politiques des gouvernements en place, mais il doit aussi anticiper les problématiques à venir et les mettre en lumière pour pouvoir les travailler aussi tôt que possible. Le cycle de trois ans, demandé par les exigences d'allers et retours dans la construction et l'expérimentation des tests préalablement à leur réalisation, nécessite de se projeter et de travailler dans une démarche prospective sur les compétences clés. Les processus ont été créés pour, et PISA a été inscrit dans une dynamique plus large d'anticipation des besoins à plusieurs décennies.

« Nous faisons évoluer PISA pour coller aux évolution d'un monde en perpétuel changement, c'est une question très importante. Ce que nous faisons, c'est que nous avons en fait trois processus en parallèle. Le premier est celui que vous connaissez, le chemin vers l'évaluation 2018. Nous préparons les tests, le cadre et tout ça. Le second est le chemin vers l'évaluation 2021. Et cela consiste à penser à comment nous allons conceptualiser ce que nous allons mesurer. En 2018 nous regardons les compétences globales, et en 2021 nous voulons regarder la créativité. Ce sont deux choses séparées. Au-delà de ça, nous sommes dans un processus que nous appelons Education 2030, dans lequel nous analysons très attentivement comment la demande de connaissances et de compétences évolue, ce que la société requiert. Nous vivons dans un monde de plus en plus incertain, de plus en plus volatile. L'idée de compétence globale par exemple est sortie en 2012, quand on a commencé à réaliser que dans le monde d'aujourd'hui, les gens peuvent voir les problèmes de différentes perspectives, avec différentes lunettes, différentes valeurs, différentes cultures. En 2012 ça sort, en 2018 on le met en pratique. C'est du long terme mais personne ne voit jamais ça, tout le monde regarde uniquement les résultats. Pourtant on passe beaucoup de temps et d'effort à penser les choses sur le long terme. »²⁸⁹

L'OCDE devient avec PISA non seulement un entrepreneur de problème public, mais il occupe également la position de pourvoyeur de « faits », donnant aux indicateurs qu'ils publient un pouvoir argumentaire fort lorsqu'il s'agit de leur faire parler de problèmes publics sur lesquels les politiques éducatives ont une possibilité d'agir. Comme le suggère Howard Becker à propos de la constitution du problème public de la consommation de marijuana dans les Etats-

²⁸⁹ *“We make PISA evolve to stick to this evolution of the changing world. It's a very important question. We basically have three parallel processes. One is you know, we are walking to the 2018 assessment. Preparing all the tests, the framework and everything. The other is, we walk to the 2021 assessment. And that's basically thinking about how we conceptualize those kind of metrics. In 2018 we want to look at global competencies, for 2021 we want to look at creativity. That's a separate thinking process further in the future. And then beyond that, we have a process called Education 2030, where we analyze very carefully how the demand for knowledge and skills and competencies is evolving, what societies require. We live in an increasingly uncertain world, an increasingly volatile world. The idea of global competencies for example came about in 2012, when people started to realize that in today's world, people can see problems through different perspectives, through different lenses, with different values, different cultures. In 2012 came the idea, and in 2018 we put it into things. It's long term but the public never sees that, they only look at the results. But we spend a lot of time and effort to things through the longer terms »* Entretien avec Andreas Schleicher, Directeur de la Direction de l'éducation, OCDE, 16 janvier 2017, traduit par nous.

Unis des années 1960, « *La plupart des citoyens comme des chercheurs en sciences sociales sentent le besoin d'étayer leurs assertions quant à l'existence d'un problème social en se référant à des faits. Ils sont sensibles à ce que des arguments dont on pourrait montrer qu'ils n'ont pas de soubassement factuel pourraient être disqualifiés facilement par des opposants.* »²⁹⁰ Les indicateurs statistiques produits par l'équipe PISA-OCDE à partir des données traitées fournies par le consortium jouent ici ce rôle de « faits », renforcé par à la fois la réputation de grande compétence des membres de l'équipe et par le traitement médiatique dont ils font l'objet au niveau international. Si les médias français ont pu, au début de l'évaluation, émettre des doutes sur l'adaptation de l'évaluation au système français dans un élan de justification des résultats décevants de la France au test - et pas forcément des élèves français, nous y reviendrons-, les dernières éditions ont vu ces objections remplacées par un discours beaucoup plus unanime sur le pouvoir explicatif des résultats. « Pisa: les inégalités entre élèves restent élevées mais stables » titrait *Libération* le 3 décembre 2019, « L'école est plus inégalitaire en France que dans la plupart des pays développés (classement PISA) » mettait en avant *France Inter* le 23 octobre 2018, « Classement PISA: La France championne des inégalités scolaires » nous disait encore *Le Monde* presque 6 ans plus tôt le 3 décembre 2013. Pas d'hypothèse dans la formulation de ces titres ni dans le contenu des articles qui les accompagnent, les données formulées dans les publications de l'OCDE sont reprises telles quelles.

Il est intéressant de noter que les publications de l'OCDE font l'objet de nombreux articles de presse au-delà de la communication ministériel qui les suit, tandis que les évaluations menées par la DEPP sur des sujets très similaires sont bien moins reprises dans les médias et la communication politique. On peut regarder le contenu des notes de la DEPP, publications classiques de la direction depuis des décennies et qui présentent avec un œil analytique les résultats des enquêtes et évaluations qu'elle mène au long de l'année. Elles ont pu par exemple

²⁹⁰ Howard Becker, *Outsiders : Studies in sociology of deviance*, Free Press, 1966, p.6, cité par Erik Neveu, La constitution d'une tradition de recherche aux Etats-Unis, Chapitre 1, dans *Sociologie politique des problèmes publics*, Armand Colin, p.30.

traiter de sujets abordés par PISA et ayant des implications sur le fonctionnement du système éducatif français, comme le redoublement ou le renforcement des inégalités au long de la scolarité, sans que cela ne soit repris par la communication ministériel ou suivi de changement d'orientation de politique éducative. Et cela bien avant que l'OCDE ne le traite dans son évaluation, qui elle a bien été utilisée par les ministres en place, parfois pour justifier des réformes d'ampleur du système éducatif.²⁹¹ Nous reviendrons sur ces usages différenciés à des fins de politique éducative par la suite. Les équipes de la DEPP, malgré le sérieux de leurs évaluations et les publications nombreuses qu'elles suscitent en interne, ne se voient pas réellement reconnaître, ni par leur ministère de tutelle ni par la Dgesco censée impulser les évolutions de politique éducative, un rôle ni d'entrepreneur de problème public, ni de pourvoyeur de « faits », du moins pas avec la même force que les équipes de PISA-OCDE.

3) Une évaluation PISA dans les faits largement orientée par l'OCDE

a) Des enjeux politiques qui amenuisent le pouvoir d'influence de la DEPP auprès de l'OCDE

Les liens entre l'équipe de la DEPP et leurs gouvernements successifs sont en fait très dépendants des personnes en poste, plus que d'une doctrine qui aurait été élaborée au lancement de l'évaluation pour garantir un lien constant entre le politique et le technique. On l'a vu, le Ministre Allègre n'avait pas renforcé l'autonomie de la DEPP dans ses travaux loin de là. Dans la même dynamique, le candidat à la présidence Nicolas Sarkozy n'avait pas fait montre d'une confiance forte dans la capacité de la DEPP à fournir des éléments de pilotage de la politique éducative par les résultats en déclarant publiquement en 2006 à propos des statistiques du ministère de l'éducation nationale :

²⁹¹ Entretien avec Bruno Trosseille, Directeur du bureau des évaluations des élèves, DEPP, 14 avril 2016.

« En inventant ses propres statistiques, une idéologie folle s'est dotée des outils indispensables à l'autocélébration de son manque de résultats. Je vous propose que nous ne connaissions qu'un seul type de statistiques : celles auxquelles font référence tous les autres pays développés. Elles seules nous disent précisément là où en est la France dans la performance de son système d'enseignement. »²⁹²

Si PISA semble être cité en filigrane comme un outil de pilotage central pour le futur Président de la République, il n'en est pas de même pour les évaluations menées par la DEPP et nulle part ne transparait son rôle d'expert français pour PISA.

A l'inverse, les liens de confiance ont pu parfois se tisser entre certains dirigeants de la DEPP et les ministres en poste lors de leur passage à la tête de la direction, et ont rendu les équipes mieux à même de faire le lien entre l'OCDE et le gouvernement. Catherine Moisan, directrice de la DEPP de janvier 2013 à novembre 2015, agrégée de mathématiques, et ancienne enseignante dans le secondaire puis à l'École normale de Rouen, avait été par le passé conseillère au cabinet d'Alain Savary ministre de l'éducation nationale, puis était entrée au cabinet de Lionel Jospin puis de Jack Lang lors de leurs passages au ministère de l'éducation nationale. Ses rencontres régulières qu'elle avait la possibilité de provoquer avec Vincent Peillon et son cabinet lors de la mise en place du cycle PISA 2015 ont permis de transmettre au politique les conclusions tirées par la DEPP à partir des résultats de la France reçues cette fois avec intérêt²⁹³. Cette opportunité presque unique de proximité entre la DEPP et le politique autour du travail sur l'évaluation PISA s'est toutefois heurtée au fonctionnement de l'administration du Ministère et notamment à l'absence de travail conjoint avec la Dgesco, pourtant en charge d'accompagner l'évolution des politiques éducatives françaises à partir des éléments de diagnostic notamment fournis par la DEPP.

²⁹² Discours d'Angers, le 1er décembre 2006, cité par Jean-Claude Emin, « Réflexions sur le rôle de la Direction de l'évaluation et de la prospective (DEP) en France », *Éducation et sociétés*, 2, n° 30, 2012.

²⁹³ Entretien avec Catherine Moisan, Directrice de la DEPP de 2013 à 2015, 10 juin 2020.

Ce flou entourant le rôle de la DEPP dans le pilotage de la politique éducative et le rôle mouvant qui lui est attribué par les décideurs politiques dans les relations avec l'OCDE a transparu dans l'ensemble des entretiens menés avec les National Project Managers rencontrés. Ils reconnaissent que le lien avec le cabinet du ministre en place est ténu, et que la décision d'intervenir ou pas auprès des équipes PISA OCDE pour faire entendre un point de vue divergent de «la France» revient davantage au directeur de la DEPP via son NPM qu'à un interlocuteur qui viendrait du cabinet du Ministre. L'un d'entre eux précisait ainsi :

«C'est le ministre qui désigne le NPM pour la France. Après il coordonne, il est l'interlocuteur de l'OCDE. C'est pas moi qui prends les décisions au final, mais c'est moi qui participais aux réunions. C'est un rouage très important, très important. Quand on n'était pas d'accord sur certaines choses, la façon d'évaluer les sciences par exemple, on prévenait le directeur, pour voir comment on pouvait intervenir. »²⁹⁴

Non seulement le NPM n'échange pas de façon organisée avec le cabinet sur les messages à porter lors des réunions internationales de PISA, mais c'est également à la DEPP de compenser les éventuelles interférences et problèmes diplomatiques engendrés par des interventions maladroites des ministres. Plusieurs situations de communication ministérielle inappropriées ont été rapportées par nos interlocuteurs, que cela concerne le contenu des publications de l'OCDE encore sous embargo, ou des publications d'autres administrations internes au ministère utilisant des données pourtant identifiées comme non publiques. Parmi ces moments délicats, on peut citer par exemple l'intervention in extremis de la DEPP auprès de la Dgescio après PISA 2008 lors de la mise en œuvre des nouveaux programmes au niveau national, pour retirer des documents d'accompagnement des programmes de mathématiques des items non libérés, ou le travail de diplomatie nécessaire lorsque les résultats de la France à PISA 2012

²⁹⁴ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 21 mars 2019.

ont été diffusés avant la fin de l’embargo imposé par l’OCDE.

« La Dgesco, ce sont eux qui font les documents d’accompagnement des programmes en mathématiques. Je ne sais pas si vous avez mis le nez dedans, mais presque tous prenaient des exemples de PISA 2008. On a même eu un problème avec ces documents parce qu’ils utilisaient des items PISA qui n’étaient pas libérés. Alors bon, on a dû les faire retirer très très rapidement, avant qu’ils soient diffusés ! »²⁹⁵

« Normalement jusqu’au mois précédent le lancement de la publication des résultats PISA, les données sont secrètes, y compris pour les cabinets. Et là, on a une journaliste de l’Express qui dévoile, bien avant la fin de l’embargo, que la France avait perdu trois places. Il y a donc eu des fuites ! Et Vincent Peillon reprend l’info à la télé. La sanction est tombée, on a eu l’interdiction d’accès à la plateforme d’échange de données. On a quand même du faire en sorte que ça n’arrive plus, ça a été compliqué. »²⁹⁶

Cet exemple, s’il montre bien l’action quasi solitaire de la DEPP dans l’accompagnement et le suivi des évaluations PISA, même une fois quelques cycles passés, est également une bonne illustration de l’absence de lien entre la DEPP co-conceptrice des épreuves PISA et réceptrice des résultats, les cabinets ministériels s’en inspirant, et la Dgesco pourtant en charge de matérialiser les changements d’orientation de politique éducative auprès des inspecteurs et des enseignants. Cette absence de travail commun a été relevée par l’ensemble de nos interlocuteurs qu’ils soient de la DEPP, de la Dgesco ou d’anciens membres de cabinets ministériels en poste entre 2000 et 2018. Elle a contribué à redéfinir le rôle de l’équipe PISA-DEPP en lui attribuant de fait une expertise technique dont elle est la seule détentrice, et une autonomie dans des choix relevant de la politique éducative française dans ses échanges avec l’OCDE et ses homologues internationaux, sans pour autant l’inclure dans un processus complet de construction de politique éducative. Un expert en somme à qui l’on confierait un

²⁹⁵ Entretien avec Franck Salles, Chargé d’étude sur l’évaluation des élèves, National Project Manager, DEPP, 11 juillet 2017.

²⁹⁶ Entretien avec Bruno Trosseille, Directeur du bureau des évaluations des élèves, DEPP, 5 janvier 2016.

travail d'expertise, mais non convié à en faire part aux moments-clés de prise de décision d'orientation de politique publique.

Ce soutien finalement faible des équipes ministérielles se double d'un processus de décision qui, lors du choix final des modalités des évaluations, semble hors de contrôle des équipes de la DEPP en charge de PISA. Un NPM partageait ainsi sa vision du processus décisionnel :

« Les décisions, c'est l'OCDE qui les prends. D'un côté tous les pays de l'OCDE, ambassadeurs des pays, décident avec les ministères de l'éducation ce à quoi ils participent. Ils nous demandent ensuite combien de modules on prend, combien on ne prend pas. Le directeur de la DEPP il ne sait pas non plus, il demande au NPM si c'est bien ou pas. Les papiers sont signés par le ministre, proposés par le directeur suite à des recos des NPM. (...) On voit bien comment ça se joue. Les politiques vont avoir une opinion, mais quand ils se retrouvent face à un technicien qui a une possibilité de décider, dans l'équipe de l'OCDE, il va dire « vous avez raison, mais techniquement c'est pas possible », ou « c'est possible mais ça va coûter tant de millions ». (...) La décision à la fin arrive au responsable de l'OCDE. »²⁹⁷

Le choix finalement est davantage celui de participer ou pas à ce qui est proposé, plus qu'à proposer de nouvelles dimensions d'évaluation ou des formules d'items différentes. L'argument politique est dépassé par l'argument technique qui peut lui même être mis à mal par l'argument financier.

Si cela est vécu parfois comme un dessaisissement des experts nationaux dans la conception des évaluations, pour les équipes de PISA-OCDE ce procédé n'est pas forcément synonyme d'une mainmise de l'organisation sur les tests. Le fonctionnement de PISA, avec ces évaluations socles doublées d'options que les pays peuvent choisir ou non, n'est pas présenté comme un moyen d'orienter le résultat final mais bien comme une façon de répondre à la diversité à la fois des contextes et des préoccupations de politique éducative des pays participants. La liberté des pays résiderait dans le choix qu'ils font de saisir telle option ou non.

²⁹⁷ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 18 mars 2019.

« L'OCDE donne les orientations, oui et non. Elle offre un menu de possibilités dans lequel les pays peuvent piocher », nous dit un ancien membre de l'équipe PISA-OCDE. Et il ajoute : « Beaucoup de propositions viennent de discussions entre les pays. L'OCDE les récupère pour ensuite les travailler dans le cadre de PISA, pour les cycles à venir. »²⁹⁸

En pratique, le choix de prendre une option ou non en est un qui porte à conséquence, puisqu'à chaque cycle les choix des pays sont observés à la fois par l'OCDE, par les autres pays participants, par les évaluateurs et les observateurs nationaux au regard politique. Parmi eux, les oppositions aux gouvernements en place bien sûr mais aussi les médias ou encore les représentants des corps d'enseignement. Par ses réunions impliquant des décisions collectives prises devant tout le monde, l'organisation internationale crée un cadre d'application d'un *soft power*²⁹⁹ puissant et sort de son rôle de simple « forum de consultation³⁰⁰ » dans lequel s'exprimeraient des experts aux avis divers cherchant à aboutir à un consensus. Et cela sans la publication ou la diffusion en interne de comptes-rendus détaillés des échanges.

Cette obligation de participation se ressent particulièrement dans PISA par rapport à d'autres évaluations menées par la même direction de l'éducation, y compris les plus visibles comme *Regards sur l'éducation*. Si la France peut y retirer régulièrement des données une fois les indicateurs produits par les équipes de EDU, matérialisées par le « w » de « *withdraw* », « retirer » en anglais, ça n'a quasiment jamais été le cas dans PISA. Le pouvoir de retirer ses données est à la fois un atout fort lorsqu'il est utilisé extrêmement ponctuellement pour faire preuve d'un réel désaccord sur le fond ou qu'il est utilisé conjointement par plusieurs pays de poids pour signifier une opposition à la démarche de l'organisation qui n'aurait pas tenu compte de critiques partagées. Un ancien membre de EDU nous indiquait ainsi la pression permise par ce « w » :

²⁹⁸ Entretien avec un agent analyste de la Direction de l'éducation, OCDE, 2 février 2017.

²⁹⁹ Le concept est inventé par le géopolitologue américain Joseph Nye en 1990, défini comme « l'habileté à séduire et à attirer », à propos de l'affirmation de la puissance dans le cadre de relations interétatiques, non fondé sur la puissance traditionnelle, la « puissance dure », exprimée dans le cadre d'un rapport de rivalité et de négociation.

³⁰⁰ Vincent Gayon, *thèse cit.*, p.236, à propos du rôle de l'OCDE dans la réalisation du rapport McCracken,

« Les pays font énormément de lobbying sur certains sujets, ils peuvent faire pression, retirer leurs données ». Que représente le fameux « withdraw » de la France sur certaines données ? Ils sont intouchables sur ces sujets. »³⁰¹

En pratique, cela n'a été utilisé qu'une fois malgré pourtant les nombreuses réserves exprimées par les évaluateurs de la DEPP, en particulier au début de l'évaluation, sur à la fois la méthode et les sujets abordés. Le travail en amont est tellement jalonné d'aller-retours avec les pays qu'un retrait de données est réellement signifiant sur le désaccord politique de la façon dont l'indicateur a été construit et les données recueillies. Un désaccord politique, alors même que finalement le politique n'a que très peu d'intervention lors de ce même processus de construction d'indicateur, on l'a vu. Le pouvoir du « w » ne semble finalement pas si fort pour les évaluateurs nationaux dans le cas de PISA. La stratégie a davantage consisté à choisir les objets d'intervention pour travailler la représentativité des données et construire un cadre d'évaluation le plus favorable possible compte tenu de la constitution du système scolaire français – ou en tout cas selon certains, le moins défavorable.

Un véritable travail a été mené sur l'intégration du concept de « grade modal »³⁰², permettant de gommer la particularité française faisant que les élèves de 15 ans « à l'heure » pouvaient se trouver aussi bien en dernière année de collège qu'en première année de lycée selon leur mois de naissance et l'âge d'entrée dans la scolarité obligatoire. Cette différence de niveau scolaire dans un système où l'apprentissage est fondé sur un programme aurait engendré un vrai écart de résultat entre les élèves de 15 ans en classe de seconde, avec donc une année de contenu en plus, et ceux en 3^e voire en 4^e s'ils avaient redoublé. La pratique du redoublement étant, au début de l'évaluation PISA, très répandue en France, l'introduction d'un grade modal a donc consisté à gommer les situations

³⁰¹ Entretien avec une statisticienne de la Direction de l'éducation de l'OCDE, 19 juin 2017.

³⁰² Felouzis Georges, Charmillot Samuel, « Comment sont construites les enquêtes PISA ? », dans : Georges Felouzis éd., *Les enquêtes PISA*, Presses Universitaires de France, « Que sais-je ? », 2012.

individuelles des élèves sortant de la « moyenne » des situations des élèves participant au test. L'OCDE en donne la définition suivante :

« Le grade modal CITE est défini ici comme le grade atteint par au moins un tiers de l'échantillon PISA. Les élèves PISA sont échantillonnés pour représenter tous les élèves de 15 ans, quel que soit le type d'établissement qu'ils fréquentent. Ils peuvent ne pas être représentatifs de leurs établissements. Réduire l'échantillon aux écoles avec le grade modal CITE des élèves de 15 ans assure que les caractéristiques des élèves choisis pour PISA représentent le profil de l'élève typique de son établissement. Le grade modal peut être celui du secondaire inférieur (CITE niveau 2) ou secondaire supérieur (CITE niveau 3), ou les deux (comme en Albanie, Argentine, Baku Azerbaïdjan, Pékin, Shanghai, Jiangsu et Zhejiang (Chine), Biélorussie, Colombie, Costa Rica, République Tchèque, République Dominicaine, Indonésie, Irlande, Kazakhstan, Luxembourg, Macao (Chine), Maroc, République Slovaque, Taipei et Uruguay). »³⁰³

Les cycles PISA les plus récents incluent donc ce grade modal dans certains de leurs tableaux. Pour ceux qui ne l'incluent pas, la France a donc fait le choix de retirer ses données, et est le seul pays à l'avoir fait. Ces interstices d'influence saisis par les évaluateurs nationaux contribuent à les rendre davantage acteurs de PISA, même si l'organisation des cycles et la répartition des tâches entre de nombreux acteurs à toutes les étapes du processus permettent à l'OCDE d'en conserver la maîtrise globale.

³⁰³ OCDE, PISA 2018 Results (Volume II) : Where All Students Can Succeed, Annexe 3, Technical notes on analyses in this volume, p. 247 : « *The “modal ISCED level” is defined here as the level attended by at least one-third of the PISA sample. As PISA students are sampled to represent all 15-year-old students, whatever type of schools they are enrolled in, they may not be representative of their schools. Restricting the sampling to schools with the modal ISCED level for 15-year-old students ensures that the characteristics of students sampled for PISA represent the profile of the typical student attending the school. Modal grade may be either lower secondary (ISCED level 2), either upper secondary (ISCED level 3), or both (as in Albania, Argentina, Baku [Azerbaijan], Beijing, Shanghai, Jiangsu and Zhejiang [China], Belarus, Colombia, Costa Rica, the Czech Republic, the Dominican Republic, Indonesia, Ireland, Kazakhstan, Luxembourg, Macao [China], Morocco, the Slovak Republic, Chinese Taipei and Uruguay)* ».

b) La non-maîtrise de construction et de la communication des résultats par les évaluateurs nationaux : un facteur d'influence majeur pour l'OCDE auprès des pays participants

i. La maîtrise de la construction des données des items aux échelles de représentation des résultats

L'organisation internationale communique sur la valeur de son processus de co-construction des cycles PISA avec les pays participants, d'innombrables experts aussi bien des disciplines que des outils statistiques faisant la force méthodologique de l'évaluation. On l'a vu, la co-construction d'une étude comprenant autant de participants et de cas particulier nécessite de trouver le plus petit dénominateur commun aussi bien sur le fond que sur la forme avec tous les compromis que cela implique. Néanmoins, le cœur de l'évaluation que constituent les items posés aux élèves, était au départ relativement dans les mains des pays chargés d'en proposer pour environ 80% du total des items retenus – les 20% étant produits par le consortium disposant également d'un groupe d'experts en interne. La doctrine de l'OCDE entre les trois premiers cycles et les suivants a changé.

Au départ, la plupart des items ont été produits par les pays participants et les groupes d'experts nationaux. En accord avec le consortium, l'OCDE avait pris le parti de demander aux pays de proposer des items à discuter par les NPM à la lumière d'échange avec des experts nationaux, avant validation collective. A partir de 2012, la conception des items a été sous traitée majoritairement à des plus petits groupes de concepteurs. Constitués de membres de plusieurs nationalités, généralement des universitaires déjà impliqués dans la mise en œuvre du test notamment en Australie, au Luxembourg, ces groupes ont conçus cette fois 80% des items, laissant 20% d'entre eux seulement à l'initiative des pays participant à PISA³⁰⁴. Le processus n'a plus fait intervenir les NPM qu'en bout

³⁰⁴ Entretien avec Franck Salles, Chargé d'étude sur l'évaluation des élèves, National Project Manager, DEPP, 11 juillet 2017.

de chaîne. Le consortium donne aux petits groupes de concepteurs la méthodologie de conception. Les items produits leur sont ensuite soumis pour première lecture, afin de vérifier qu'ils correspondent bien aux parties du cadre d'évaluation qu'ils sont censés mesurer. C'est une fois ce tri effectué que les items sont présentés aux NPM pour qu'ils puissent vérifier qu'il n'y a pas d'erreur technique oubliée, et surtout qu'il n'y a pas de risque de biais culturel. Les items étant ancrés dans des situations de la vie réelle, l'importance du contexte est primordiale dans la façon dont les élèves peuvent répondre, au-delà de leurs compétences dans la discipline évaluée. Finalement ce ne sont plus que les biais inacceptables qui ont motivé le retrait d'items, alors qu'à l'origine la force de proposition émanait des NPM.

Ce morcellement des compétences s'est accru de cycle en cycle. Les pôles de gestion des tests PISA ont été divisés entre plusieurs acteurs à partir du cycle 2015, et le pôle de coordination centrale des évaluations a notamment été supprimé pour être repris en main par l'OCDE. Le consortium était en effet jusqu'alors organisé en plusieurs pôles dont celui de coordination, occupé depuis 2000 par ACER, en plus de la tâche d'organiser la passation des tests sur le terrain et la correction des réponses reçues. A partir du cycle 2015, les rôles ont été différemment répartis entre Pearson, chargé de structurer le cadre et les orientations données par l'OCDE à partir des conclusions des réunions des pays participants, et ETS travaillant pour sa part sur le cœur de l'évaluation, la conception de la majeure partie des items, la conception des analyses psychométriques, la mise en œuvre des tests et la correction des tests passés. Cette nouvelle organisation a engendré une moindre implication des NPM dans les prises de décision, sans que leurs remarques sur le besoin d'être davantage sollicités ne soient réellement transmises au CPP. Si le portage de ce message a bien été fait par la France par la voix du représentant présent au CPP, Thierry Rocher, cela n'a pas forcément été le cas dans d'autres pays dans lesquels les deux protagonistes nationaux ont pu être moins proches³⁰⁵.

³⁰⁵ Entretien avec Franck Salles, Chargé d'étude sur l'évaluation des élèves, National Project Manager, DEPP, 11 juillet 2017.

La seconde partie du processus de traitement des données rend cette fois quasi impossible la compréhension fine du mode de calcul puisque les données « brutes » fournies aux pays représentant les résultats de leurs élèves ont déjà été pondérées par le consortium. Les statisticiens de l'OCDE coordonnent la constitution des bases de données sur lesquelles ils fonderont leurs analyses et informent les pays de l'avancée des travaux. Les NPM quant à eux collectent les données mais elles sont intégralement traitées et redressées par ETS ou ACER avant cela. Les données qui leur sont fournies pour relecture sont donc bien différentes, sans que la façon dont elles aient été obtenues soit très clairement expliquées ni même reproductible. Cette « boîte noire ³⁰⁶ » pour les évaluateurs nationaux constitue une vraie difficulté qui complique non seulement la tâche de vérification et d'analyse mais aussi celle de communication voire de pédagogie autour des publications lors de leur sortie.

ii. Une maîtrise du tempo de la communication et donc de l'analyse des résultats par les évaluateurs nationaux

La maîtrise de la communication qui entoure la publication des résultats de PISA s'est construite depuis le départ comme un enjeu fort du Programme international³⁰⁷. Cette dimension un peu à part de l'évaluation, qui arrive en fin de cycle une fois que les résultats sont sortis et ne peuvent donc plus être travaillés par les pays dans le cadre de la publication de l'OCDE, est tout à fait fondamentale pour comprendre la posture d'expertise de l'équipe PISA-OCDE et les frictions qu'elle peut générer avec l'équipe de la DEPP. Enjeu stratégique, elle se veut en droite ligne des principes de la communication politique en général, se dessinant autour des positions de pouvoir des différents acteurs qu'elle fait intervenir³⁰⁸. Loin de n'avoir qu'une fonction instrumentale d'outil visant à convaincre ses récepteurs

³⁰⁶ Entretien avec une statisticienne, en charge de PISA, DEPP, 1er juillet 2020.

³⁰⁷ Alain P. Michel, Nathalie Mons, « PISA et les politiques éducatives : une influence complexe », *Administration & Education*, 1, n°145, 2015.

³⁰⁸ Jacques Gerstlé, *La communication politique*, Armand Colin, Paris, ed. 2011.

Les consignes sont claires. D'un côté, les règles d'embargo sont strictes. Aucun résultat, aucun indicateur, aucun classement, même s'il est évidemment connu des évaluateurs nationaux avant la sortie des volumes puisqu'ils en ont à plusieurs moments la relecture, ne doit être divulgué avant la publication officielle dont la date est décidée par l'OCDE et dont la présentation dans plusieurs pays du monde est faite par des agents de l'OCDE, qu'ils soient membres de l'équipe PISA, à la tête de la direction toute entière voire membres du cabinet du Secrétariat général. D'un autre côté, l'exigence de transparence et de probité qui entoure PISA a incité l'OCDE avec l'accord des pays participants à rendre une majeure partie des travaux publics. Les volumes publiés se veulent les plus exhaustifs possibles pour donner à voir non seulement les résultats en eux-mêmes mais aussi les mécanismes de construction des indicateurs publiés. Un volume est notamment dédié à l'explication presque in extenso de la méthodologie employée pour chaque cycle, et est laissé en libre accès sur la bibliothèque en ligne de l'organisation. C'est bien ce « presque » qui a toute son importance dans la maîtrise de l'usage des données publiées finalement conservée par l'OCDE.

La compréhension de la méthode – ou plutôt sa non-compréhension par les lecteurs nationaux, soit par manque de compétences techniques soit par l'opacité du processus- favorise l'interprétation des résultats avec une logique de classement brut, sans capacité de comprendre ni les causes des phénomènes observés ni les conclusions possibles en termes de politique éducative.

« C'est difficile de saisir clairement la signification des résultats. Il y a un problème, on peut avoir les scores réels, mais pas les scores logarithmiques. On s'aperçoit que sur 70, cinq ou six pays ont 59. On ne peut pas les classer. Ceux qui ont 60 aussi. Donc on n'arrive pas à voir de différence. En prenant une autre échelle qui fait des zooms on peut avoir des différences. Une fois qu'on a fait ça on ne peut plus revenir en arrière. Même s'ils ont un peu changé la méthode, ils sont obligés de faire comme ça. Et du coup même si on ne met pas les scores sous forme de liste, les journalistes peuvent le faire. Ils vont dire c'est pas nous mais ça permet de faire ça. »

« Ils ne savent pas comment est faite l'échelle. Les journalistes, les enseignants, les politiques. Les gens ne savent pas ce que c'est 270. Même un inspecteur d'académie. Peu de gens savent comment ça se passe. »³⁰⁹

Finalement, c'est la logique de classement qui prédomine dans les éléments de langage non seulement repris par les médias et les décideurs politiques –en particulier les ministres de l'éducation en place à chaque sortie de volume -, mais aussi fournis par l'OCDE.

Les cycles de trois fois trois ans permettraient à chacun de comparer l'évolution de ses résultats sur presque une décennie, soit la quasi-totalité de la durée de l'éducation obligatoire pour un pays comme la France – à l'époque de 6 à 16 ans. Pourtant la mise en avant du pouvoir de classement est plus ou moins explicitement maniée dans les publications PISA. La dernière en date, dans son volume 1, commence même sa préface générale par ces mots :

« Il ressort entre autres innombrables constats de l'évaluation PISA de 2018 que les élèves de 15 ans des quatre provinces/municipalités chinoises participantes – Pékin, Shanghai, Jiangsu et Zhejiang – l'emportent haut la main en mathématiques et en sciences sur ceux des 78 autres systèmes d'éducation participants. »³¹⁰

C'est un véritable palmarès qui est alors fait entre tous les participants à PISA 2018. L'objectif qu'attribuait Schleicher à PISA à son origine, celui de servir de miroir utilisable dans n'importe quel contexte national dans le monde, s'est mué en une course au meilleur résultat possible porté au pinacle par des publications de l'OCDE mondialement lues. Finalement, l'important ne semble plus réellement de mesurer ce que à quoi on accorde de la valeur en matière de politique éducative – dans le cas de PISA, permettre aux élèves d'avoir les meilleures performances scolaires possibles – mais plutôt l'inverse : on accorde de la valeur à ce que l'on évalue. Les espaces d'échanges sur les projets de politiques éducatives de chacun et les retours d'expérience constructifs pour améliorer le pilotage de son action publique paraissent n'être plus qu'en arrière-

³⁰⁹ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 18 mars 2019.

³¹⁰ OCDE, *PISA 2018 Volume 1*, Editions OCDE, 2019.

plan.

Cette communication sous forme de classement faite par l'OCDE des résultats des pays dissone de la vision portée par le mythe créateur de PISA. Elle dissone également des principes qui ont régi officiellement le processus de création de l'évaluation.

Pour les évaluateurs français, tout l'enjeu a consisté d'abord à convaincre de l'intérêt de raisonner en comparant dans le temps les résultats de la France par rapport à elle-même, et non pas par rapport aux autres pays des classements. Elle a ainsi communiqué sur les résultats français par de multiples notes et publications auprès de tous ceux susceptibles de devenir les utilisateurs des données publiées, à savoir les décideurs politiques et dans une moindre mesure les chercheurs et pairs évaluateurs³¹¹. Il faut noter que peu de travaux ont été construits pour être dirigés vers les corps de l'Education Nationale, les inspections académiques notamment. Si ce sujet n'est pas l'objet de cette thèse, il est toutefois intéressant de remarquer, sans pouvoir y apporter de réponses, que PISA n'a pas été conçu ni exploité au niveau national comme un outil à transmettre aux corps enseignants ou d'inspection de l'éducation nationale. Certains éléments issus des évaluations PISA ont pu être transmis néanmoins aux enseignants lors de formation, à la faveur de l'intervention de membres de l'équipe PISA-DEPP, enseignants à temps partiels, et impliqués pour la préparation des contenus de formation. Certains items libérés de PISA ont par ailleurs été repris dans les manuels scolaires à partir des années 2010³¹², sans que cela ne soit toutefois explicitement précisé ou que cela fasse l'objet d'un accompagnement pédagogique spécifique.

Malgré ces efforts pour orienter l'interprétation des résultats publiés par l'OCDE, la logique de classement a perduré dans les éléments de langage utilisés par les différents ministres et par les médias à chaque nouvelle sortie d'un volume

³¹¹ Toutes les publications de la DEPP ou de ses équipes sur PISA sont disponibles sur le portail Educ Eval du ministère de l'Education nationale sous l'onglet « Evaluations internationales » - PISA : <http://educ-eval.pleiade.education.fr/pisa7.htm>

³¹² Entretien avec Franck Salles, Chargé d'étude sur l'évaluation des élèves, National Project Manager, DEPP, 11 juillet 2017.

PISA. La DEPP a donc complété sa stratégie de communication à partir du cycle 2018 pour intégrer une conférence de presse préalable à la sortie du volume principal de PISA 2018, à destination exclusivement des journalistes. Cette conférence de presse a pris la forme d'une session de formation et d'information sur PISA, sa méthode, ses indicateurs, le sens à donner aux valeurs indiquées dans les différents tableaux et graphiques, afin qu'ils puissent interpréter plus facilement les résultats et construire différemment leurs articles à venir. Jugée très satisfaisante et constructive par les évaluateurs de la DEPP, cette démarche a même selon eux inspiré l'OCDE qui se serait placée dans son sillage pour l'organisation de ses conférences de presse³¹³.

Cette reprise de l'expérimentation française de format nouveau de conférence de presse est d'autant plus perçue comme importante que le moment de la publication est au cœur du rapport de force entre l'OCDE et les pays participant à PISA. L'organisation a bien la primeur de la communication à tout prix. Cette maîtrise absolue du tempo de publication est primordiale pour permettre à l'OCDE d'être la première voix à s'exprimer sur les résultats, la première voix à donner des conclusions qui seront reprises par les médias du monde entier – et qui pourront par là même éclipser certains résultats qui auraient pu être davantage mis en avant. Bien entendu, aucun résultat, même s'ils ont été relus par les évaluateurs nationaux, ne doit être communiqué avant la levée officielle de l'embargo. Les données ne sont pas à la main des pays, ni des administrations et encore moins des ministres. Des communications trop empressées ont pu engendrer de vraies tensions entre l'OCDE et les évaluateurs de la DEPP – même s'ils n'étaient en rien responsables de la « fuite » faite par leur ministre. Ce fût le cas par exemple, cité par plusieurs de nos interlocuteurs à la fois de l'OCDE et de la DEPP, des quelques jours qui ont entouré la publication des résultats de PISA 2012. Le ministre de l'époque Vincent Peillon s'était exprimé publiquement sur les résultats français dans les médias avant même la fin de l'embargo. Cette dérogation au protocole de publication des résultats avait été très peu appréciée

³¹³ Entretien avec Ronan Vourc'h, Adjoint au chef de bureau de l'évaluation des élèves, DEPP, 22 juin 2020.

par l'organisation internationale et la DEPP avait dû faire en sorte d'éviter toutes représailles dans les cycles suivants, notamment dans l'accès aux données avant la sortie officielle des publications.

La simultanéité de la communication de l'OCDE sur ce qu'elle souhaite valoriser dans chaque publication et la sortir officielle de l'intégralité des indicateurs des volumes sortis crée un véritable challenge analytique pour les évaluateurs de la DEPP qu'ils n'ont pas encore tout à fait réussi à relever. Le temps est beaucoup trop court pour pouvoir réellement construire une analyse plus poussée sur certains points potentiellement intéressants dans le cas de la France mais non explorés pour l'ensemble des pays participants par l'équipe PISA-OCDE. L'équipe de la DEPP a dû donc se résigner depuis le début à sortir une note d'information rapidement, pour pouvoir être perçue comme l'interlocuteur référent pour la France. Le titre des notes, tous sur le même modèle depuis le premier cycle 2000, est parlant : « *Elèves de 15 ans : premiers résultats de l'évaluation internationale de PISA (année) en (discipline majeure)* ». Sur un format de 4 pages, on pourrait attendre une seconde analyse plus poussée des résultats publiés par l'OCDE sous la forme de rapports de plusieurs centaines de pages. En réalité, le temps contraint à disposition des équipes a souvent rendu cette analyse trop complexe à mener. La stratégie adoptée a donc été la publication de ces résumés très proches des « country notes » de l'OCDE centrées sur chaque pays participant, des « résultats froids » comme ont pu les nommer nos interlocuteurs de la DEPP³¹⁴, suivie sur certains sujets très précis d'articles publiés par exemple dans la revue Education, formation ou la revue de l'INSEE France portrait social, ou les Dossiers du ministère de l'Éducation nationale.³¹⁵ Ce renoncement, lié à la fois à la communication de l'OCDE en amont et au moment de la publication des résultats et à l'organisation interne de la DEPP

³¹⁴ Entretien avec une statisticienne, en charge de PISA, DEPP, 1er juillet 2020.

³¹⁵ Voir par exemple Isabelle Robin et Thierry Rocher, « La compétence en lecture des jeunes de 15 ans : une comparaison internationale », *Données sociales : la société française – Edition 2002-2003*, INSEE. ; Fabrice Murat et Thierry Rocher, « La place du projet professionnel dans les inégalités de réussite scolaire à 15 ans », *France, portrait social 2002-2003*, INSEE ; DEPP, *L'évaluation internationale PISA 2003 : compétences des élèves français en mathématiques, compréhension de l'écrit et sciences*, les Dossiers Enseignement scolaire, n°180, mars 2007.

travaillant sur d'innombrables études en même temps avec une équipe réduite, contribue à relativiser l'expertise de la DEPP dans l'accompagnement des décisions de politique éducative.

Conclusion du chapitre

Au moment du lancement de PISA, la gestion publique française se trouve à un tournant majeur en se lançant dans une réforme inédite, la LOLF, lui attribuant un pilotage par les résultats. Alors que la place donnée aux indicateurs de résultats aurait dû devenir prépondérante dans la façon de piloter les politiques publiques et parmi elle la politique éducative, il n'en est finalement rien. La réflexion sur les objectifs de la politique éducative n'est finalement pas aboutie, et les indicateurs employés, du moins les premières années, s'apparentent davantage à des données descriptives du système éducatif qu'à de véritables outils de pilotage correspondant aux enjeux éducatifs d'un monde en changement de plus en plus rapide. De même, alors que les évaluateurs statisticiens de la DEPP étaient tout trouvés pour contribuer à matérialiser ce nouveau paradigme, le débat sur l'indépendance d'un service statistique interne au ministère dont il est censé évaluer l'action vient prendre le pas sur un travail de construction d'un programme de travail correspondant aux besoins d'évaluation.

Dans le même temps, ces mêmes évaluateurs sont sollicités et laissés seuls pour représenter la France auprès de l'OCDE dans le lancement de la plus grande évaluation internationale des résultats des élèves, qui devient en quelques années incontournable dans le pilotage de la politique éducative nationale. Face à des agents internationaux reconnus par tous pour leur haute expertise à la fois statistique et d'analyse des politiques éducatives, les agents de la DEPP peinent à imposer leur propre expertise malgré des profils et des parcours finalement assez similaires à ceux de leurs interlocuteurs. Finalement la ressource du pouvoir promise aux évaluateurs de la DEPP au moment de l'attribution de leurs missions dans le suivi et la co-construction de l'évaluation PISA avec les équipes

de l'OCDE ne remplit plus exactement son rôle. Les experts de la DEPP, s'ils ont réussi tant bien que mal à conserver à la fois une voix dans la façon dont sont construits les cycles PISA et l'oreille – bien que lointaine la plupart du temps – des cabinets ministériels, ont difficilement préservé leur pouvoir d'influence auprès de l'OCDE d'une part et les décideurs politiques nationaux d'autre part. Cette difficile maîtrise stratégique autour à la fois de la reconnaissance de la spécificité de leurs compétences et de la communication d'information autour des résultats français de PISA a eu un effet concret sur la façon d'organiser les relations avec les équipes PISA-OCDE et de s'approprier les résultats français à l'évaluation internationale. Les manifestations d'influence des évaluateurs français sont alors tout à fait représentatives d'un pouvoir qui n'est pas lié à une position statutaire, finalement très similaire dans les textes lorsque l'on considère nos évaluateurs nationaux et internationaux, mais qui s'inscrivent dans les relations qu'ils entretiennent les uns avec les autres³¹⁶.

Se tissent des relations entremêlées de manifestations d'un soft power fort, aussi bien sur le fond de l'évaluation (les sujets abordés) que sur la forme (la méthode statistique employée, la construction des items) et la communication des résultats.

Du côté de l'OCDE, l'implication des équipes d'évaluateurs nationaux, organisée à la fin des années 1990 avec l'accord des pays participant à PISA, entraîne finalement en sous-main des sentiments paradoxaux au sein des agents de la DEPP: d'une part une obligation de contribution (« il faut être là »), qui refreine l'exposition des désaccords et implique une participation à une évaluation sur laquelle les marges de manœuvre vont en diminuant, et d'autre part un sentiment d'appartenance à une communauté d'experts (« je suis à ma place parmi ces experts internationaux ») qui contribue à renforcer l'expertise des agents de la DEPP.

Du côté de la DEPP, le pouvoir s'exprime dans les interstices qui lui sont laissés par l'organisation internationale et par les décideurs politiques français,

³¹⁶ Michel Crozier et Ehrard Friedberg, *L'acteur et le système*, Paris, Ed. du Seuil, 1977.

pour la plupart désengagés de la construction des cycles de PISA. Avec parcimonie, ils influent sur la méthode pour qu'elles soient la mieux adaptée au fonctionnement du système français et sont suivis par d'autres pays participants aux problématiques finalement similaires. L'intégration d'un grade modal en est un parfait exemple. Ils travaillent aussi à maîtriser davantage la communication autour des résultats des tests des élèves français, afin de lutter contre un effet « classement » souvent défavorable à la France et qui coupe court à des réflexions plus poussées sur les implications de tels résultats. Avec une équipe restreinte, des moyens sans hausse et un programme de travail de plus en plus important, difficile de prendre le temps de retravailler une telle masse de données que celles produites dans le cadre de PISA, et d'autant plus lorsque le rythme des publications choisi par l'OCDE est resserré.

Les premiers cycles de PISA n'ont pas permis à la DEPP de faire profiter leurs évaluations nationales de l'expertise et de la notoriété acquises par leur implication dans l'évaluation internationale. En réaction, le traitement des résultats est, nous le verrons, fait sous un œil particulièrement critique à la fois de la méthodologie de l'évaluation et du positionnement de l'OCDE dans la construction des cycles. La seconde période, à partir des années 2010, marque en revanche un changement de perspective. Forts d'un contexte politique plus favorable à leurs travaux, les agents de la DEPP construiront leurs propres outils cette fois bien identifiés par les décideurs, un savant mélange des évaluations nationales fondées sur les programmes et des principes novateurs de PISA, avec l'évaluation CEDRE.

Chapitre 4 : La cristallisation des enjeux d'influence à travers l'usage de l'outil.

A l'arrivée de PISA dans son programme de travail, la DEPP devient l'opérateur français unique de l'évaluation internationale. Au-delà de ce rôle de coordination et de suivi de l'organisation de l'évaluation pour les élèves et les établissements scolaires français, elle se donne également pour tâche de fournir une description analytique des résultats de la France à l'évaluation, à destination des décideurs politiques. En tant que direction au sein du ministère, et même si elle ne construit pas les nouvelles orientations de politique éducative voulues par les ministres successifs, elle conçoit son action comme relevant de l'apport d'une expertise forte aussi bien du point-de-vue technique que de celui de la compréhension du fonctionnement du système éducatif français à travers les résultats des évaluations qu'elle mène.

Alors que l'on pouvait attendre une prise en main des résultats de PISA par la DEPP dès le départ, compte tenu d'une part de cette volonté de la DEPP de valoriser son expertise en matière d'évaluation de la politique éducative et d'autre part de l'investissement financier que l'évaluation de l'OCDE a nécessité en interne au ministère de l'Education nationale, elle semble cependant réduite au strict minimum. Quelques notes d'information très synthétiques sont publiées, mais la DEPP ne s'engage pas davantage dans la valorisation de travaux auxquels elle a pourtant participé depuis la conception jusqu'au lancement sur le terrain des tests et à la validation des données recueillies.

Nombreux sont les travaux de recherche et articles de presse à mettre en avant l'absence d'un « choc PISA » français³¹⁷, comme il aurait eu lieu dans d'autres

³¹⁷ Voir par exemple les travaux de Nathalie Mons et Xavier Pons sur PISA, et notamment « Pourquoi n'y a-t-il pas eu de « choc PISA » en France ? Sociologie de la réception d'une enquête internationale (2001-2008) », *Revue française de pédagogie*, 182, 2013 ; *La réception de PISA*

pays, y compris des pays européens voisins. L'expression aurait même été inventée pour parler très spécifiquement du cas de l'Allemagne³¹⁸, citée régulièrement en exemple d'un pays ayant pris acte de ses résultats en dessous de la moyenne des pays de l'OCDE au premier cycle de PISA en 2000, pour construire dès 2001 un grand plan d'action éducatif³¹⁹. Au-delà du constat de résultats significativement inférieurs à la moyenne dans toutes les disciplines et un classement global médiocre, l'analyse des décideurs publics allemands a en particulier porté sur l'inéquité du système éducatif qu'ils lisaient à travers l'évaluation. L'écart important et inattendu entre la perception d'une grande qualité du système éducatif et les résultats décevants à l'évaluation a alors provoqué un « tollé général³²⁰ » en Allemagne. Le pays a engagé à grande vitesse des réformes structurelles importantes, et parmi elles la refonte de l'enseignement secondaire relevant pourtant de la compétence des Länder, l'organisation d'un système d'inspection des établissements, et une démarche d'évaluation comparative du système allemand avec d'autres systèmes scolaires.³²¹

En France en revanche, pas de « choc PISA » n'a été observé au sens que lui donnent Xavier Pons et Alain P. Michel d' « *évolution convergente et brutale du débat public vers l'acceptation d'un ensemble de réformes dont PISA aurait montré la nécessité* »³²², du moins pas dès le départ. Le débat public sur le sujet est en fait presque inexistant à la sortie des premiers résultats. Les références à PISA

en France. Connaissance et régulation du système éducatif, 2009, Paris, OSC. Voir aussi les travaux de Christian Baudelot et Roger Estabiet, et notamment *L'élitisme républicain : l'école française à l'épreuve des comparaisons internationales*. Paris, Ed du seuil, 2009.

³¹⁸ Dennis Niemann. « Le « choc PISA » en Allemagne et les réformes de l'éducation », *Administration & Éducation*, vol. 145, n° 1, 2015, pp. 39-44 ; Jeanne Rubner, « Les suites du choc PISA dans les médias allemands », *Revue internationale d'éducation de Sèvres*, 66, 2014.

³¹⁹ Voir par exemple Dennis Niemann. *Art. cit.* ; Denis Niemann et Kerstin Martens, « When Do Numbers Count ? The Differential Impact of Ratings and Rankings on National Education Policy in Germany and the US », *German Politics* 22 (3), 2013 ; Klaus-Jürgen Tillmann, Kathrin Dederig, Daniel Kneuer, Christian Kullmann et Isa Nessel, *PISA als bildungspolitisches Ereignis. Fallstudien in vier Bundesländern*, Wiesbaden : VS Verlag für Sozialwissenschaften; 2008.

³²⁰ Andreas Ammermüller, « PISA : What Makes the Difference ? », Explaining the gap in PISA test scores between Finland and Germany', ZEW Discussion Papers, n° 04-04, Centre for European Economic Research, Mannheim, 2004.

³²¹ Claire Dupuy, « Les politiques régionales d'éducation en Allemagne depuis PISA », *Revue internationale d'éducation de Sèvres*, n° 64, 2013.

³²² Alain P. Michel, Xavier Pons, « Éditorial. Le « choc » PISA : métaphore ou révolution ? », *Administration & Éducation*, vol. 145, n° 1, 2015.

demeurent ainsi très rares dans les discours des politiques et des directions du ministère de l'Éducation nationale, dans la presse et les travaux de recherche français.

En particulier, la communication faite par les médias est extrêmement limitée, à la fois en termes de nombre d'articles publiés mais aussi en termes de diversité de dimensions de PISA analysées par les journalistes et les commentateurs³²³. Dans un travail d'analyse de dépêches de l'AEF traitant de PISA, réalisé par Nathalie Mons et Xavier Pons³²⁴ à la fin des années 2000 et portant donc spécifiquement sur les trois premiers cycles de l'évaluation, on observe ainsi que le nombre de dépêches évoquant PISA est peu élevé entre 1999 et 2009, malgré quelques pics au moment de la sortie des publications en fin d'années 2001, 2004 et 2007. Se situant habituellement autour d'une dépêche par mois, elles peuvent atteindre une douzaine en décembre 2001, une vingtaine en décembre 2004 ou encore moins d'une quinzaine en décembre 2009. Il faut souligner qu'elles abordent en fait peu le cas français, traitant pour la plupart des résultats des autres pays ou reprenant les communications faites par l'OCDE elle-même. Du côté des titres de presse quotidiens, seuls quatre quotidiens nationaux évoquent les résultats de PISA 2000 à la fin de l'année 2001, *Le Monde*, *Le Figaro*, *La Croix* et *l'Humanité*, et l'hebdomadaire *Le Point*. Ils abordent les résultats des élèves français d'une façon relativement similaire, faisant le constat, plus ou moins sévère selon la couleur politique des rédactions, de leur médiocrité. *Le Monde* du 5 décembre 2001 parle ainsi de la France comme d'un « élève moyen », *Le Figaro* du même jour titre quant à lui : « Le système scolaire français : peut mieux faire ». *L'Humanité* du 7 décembre approfondit l'analyse de ces compétences moyennes en précisant dans son titre : « les élèves français manquent d'analyse critique, selon l'OCDE ». Mises à part ces rares reprises, le

³²³ Alain P. Michel, Nathalie Mons, « PISA et les politiques éducatives: une influence complexe », Association Française des acteurs de l'éducation, *Administration & Education*, 1, n°145, 2015.

³²⁴ Nathalie Mons, Xavier Pons, avec la participation d'Agnès Van Zanten et Jeremy Pouille, *La réception de PISA en France, Connaissances et régulation du système éducatif*, KNOWandPOL, orientation 3, avril 2009, p.9.

traitement médiatique semble faible : la conférence de presse organisée pour accompagner la publication des résultats attire peu de journalistes, pas plus que les notes d'information qui traiteront ensuite du sujet.

Cette réception particulière de PISA, ressemble à première vue à une non-réception³²⁵ : PISA ne va pas contribuer dans un premier temps à constituer les performances du système éducatif français en problème public. Cette non-réception médiatique lors de la publication des premiers résultats reflète sans doute la dynamique des échanges avec les médias : absence de communication organisée de la part des évaluateurs nationaux ou des décideurs politiques ; faible écho des travaux au sein de la DEPP dans les médias, dont les études ne sont ordinairement pas réellement publicisées ; désintérêt des journalistes français pour ce que la DEPP considère comme des problèmes méthodologiques dans la conception de l'évaluation³²⁶ et sur lesquels seront initialement centrés les échanges avec la presse.³²⁷

Mais on ne peut limiter cette non-réception de PISA, à un simple problème de communication. L'arrivée de PISA dans la boîte à outils des évaluations disponibles en France aurait pu permettre aux évaluateurs français de s'en saisir afin de disposer de nouvelles données de pilotage de l'action publique en éducation. Cependant, ce n'est pas cet usage qui est privilégié au départ. Nous verrons que cette apparente non-réception peut s'expliquer par plusieurs processus. Au-delà du difficile positionnement de la DEPP au niveau national déjà évoqué dans les chapitres précédents, elle reflète un contexte spécifique dont nous nous attacherons à décrire les points saillants, et en particulier l'expérience ratée de la participation de la France – par le biais de la DEPP – à l'évaluation IALS sur la littératie des adultes. Elle est en outre la première étape d'une

³²⁵ Voir notamment Nathalie Mons, Xavier Pons, *op. cit.*; Norbert Bottani et Pierre Vrignaud, *La France et les évaluations internationales*, Rapport national au HCEE, Paris, 2005 ; Nathalie Mons, « L'évaluation des politiques éducatives. Apports, limites et nécessaire renouvellement des enquêtes internationales sur les acquis des élèves », *Revue internationale de politique comparée*, 14, 3, 2007.

³²⁶ Ces problèmes perçus par la DEPP seront largement repris dans ce chapitre. Il s'agit principalement de l'utilisation de méthodes de test psychométriques, notamment de questionnaires à choix multiples, et des méthodes de traduction des items.

³²⁷ Nathalie Mons et Xavier Pons, *op. cit.* p. 24.

stratégie plus longue de valorisation des évaluations nationales menées par la DEPP que nous décrirons dans ce chapitre, qui commencera par un premier temps de critiques fortes et ouvertes de PISA pour s'acheminer vers une réappropriation des principes méthodologiques nouveaux appliqués dans l'évaluation internationale pour légitimer sa maîtrise des évaluations en général, et l'expertise fournie par ses propres évaluations.

Nous verrons ainsi dans un premier temps que les réactions des agents de la DEPP, et plus spécifiquement ceux impliqués dans la mise en œuvre de PISA, seront publiquement très critiques sur la validité de l'enquête concernant le système français. Bien que l'évaluation de l'OCDE soit prometteuse pour permettre à la DEPP d'asseoir publiquement son rôle d'entrepreneur de cause, ils ne trouvent pas au lancement de PISA l'outil espéré pour légitimer leur apport de connaissances dans les différentes phases de la construction des problèmes publics de politique éducative. Ramené à un rôle qu'ils décrivent comme de simple exécution, sans possibilité de valoriser leur expertise, ils perçoivent PISA comme une concurrence, non une ressource nouvelle. Dès lors, s'élabore au sein de la DEPP un cadre d'analyse critique du programme d'évaluation international permettant de revendiquer la maîtrise technique, voire le monopole, de l'évaluation et du fonctionnement du système éducatif français. La réception qui est faite publiquement par la DEPP des premiers résultats PISA se résume plus spécifiquement à la formulation de critiques portées à l'égard de la méthode de PISA et à la recherche de méthodes alternatives pour promouvoir des enquêtes standardisées correspondant mieux à leurs attendus de qualité et de représentativité du système français. Les difficultés méthodologiques qu'ils identifient, et pour lesquelles ils ont même parfois exprimé des réserves avant même la passation de premiers tests, expliqueraient pourquoi les résultats des élèves français étaient moins bons qu'attendus. Nous nous placerons dans la lignée des travaux de Nathalie Mons et de Xavier Pons qui ont notamment étudié la réception de PISA en France par tous les acteurs du système éducatif, en concevant la réception de l'évaluation non pas comme un mécanisme unilatéral de transmission d'information mais bien comme une activité faisant intervenir

les rapports d'intérêts et de force des différents destinateurs et destinataires³²⁸. Pour ces auteurs, les trois premiers cycles de PISA n'ont pas suscité de réflexion collective pour trouver des solutions à des situations perçues comme problématiques, ni dans le champ médiatique, ni dans le débat politique, ni dans le débat scientifique. Dans le cas des évaluateurs de la DEPP, la réception de PISA serait corrélée à leurs modalités et l'intensité de leur implication dans la construction ou la tenue de l'évaluation³²⁹, sur ces trois premiers cycles. En d'autres termes, leur faible appropriation des résultats est la conséquence d'une faible mobilisation des acteurs. Nous verrons que si cette hypothèse peut justifier leur regard aiguisé sur la méthode de PISA au moment de sa construction et de son lancement, elle se place uniquement du côté de la réaction à l'évaluation internationale. Or il faut, à nos yeux, aussi prendre en compte la construction en parallèle d'une stratégie de reconquête d'une légitimité que les évaluateurs nationaux jugent perdue, non seulement de leur crédibilité d'expert auprès de leurs homologues internationaux mais aussi de leur positionnement comme acteur incontournable, en tant que DEPP, des politiques éducatives en France.

Nous verrons en effet dans un deuxième temps que se manifestent derrière cette critique méthodologique, des tentatives de valorisation des évaluations nationales et de l'expertise des évaluateurs nationaux de la DEPP, grâce au travail qu'ils mènent pour PISA. Cette stratégie de légitimation de leur rôle d'expert en évaluation part du constat que la France continuera à participer à PISA, et cela malgré leurs alertes nombreuses sur les inadaptations de la méthode choisie par l'OCDE pour regarder avec la justesse statistique requise les résultats de l'évaluation du système éducatif français en comparaison de ceux des autres participants. Le choix se porte alors, par intérêt des acteurs pour la découverte de nouvelles conceptions des résultats de l'éducation comme par pragmatisme pour renforcer leur statut d'expert des évaluations en éducation, sur le lancement

³²⁸ Ils s'inscrivent dans la lignée des travaux de Eric Mangez, *Réformer les contenus d'enseignement*, 2008, Paris, PUF.

³²⁹ Nathalie Mons, Xavier Pons, avec la participation d'Agnès Van Zanten et Jeremy Pouille, *La réception de PISA en France, Connaissances et régulation du système éducatif*, KNOWandPOL, orientation 3, avril 2009, p. 6.

rapide d'un travail s'apparentant presque à un syncrétisme, visant à construire autour de PISA les outils d'évaluation nationale les plus complets et les plus visibles possibles. Dès les premières années de la décennie 2000, la DEPP lance en particulier l'évaluation CEDRE, évaluation nationale reprenant plusieurs des piliers méthodologiques de PISA et notamment la démarche psychométrique et l'intérêt pour le non-disciplinaire. Cette évaluation servira de faire-valoir pour le savoir-faire technique français en matière d'évaluation et contribuera à vérifier les résultats de PISA en appliquant une méthodologie proche mais jugée par la DEPP plus adaptée aux particularismes du système éducatif français. La mise en œuvre de CEDRE sera par ailleurs au centre d'une vaste entreprise de recentrage de PISA vers le niveau national, afin de redonner aux agents de la DEPP le contrôle sur l'analyse de ce qui peut être déduit des résultats de PISA ou ce qui est, à leurs yeux, peu représentatif de la situation française. PISA devient, dans le discours de la DEPP, une évaluation parmi d'autres, permettant, lorsque les résultats sont similaires aux leurs, de valoriser par sa visibilité et sa légitimité auprès des décideurs politiques, les évaluations « maisons », et notamment CEDRE.

Cette stratégie de long terme de la DEPP, loin de se résumer à de seules critiques de méthode ou à une opposition de modèles d'évaluation, a contribué à réaffirmer son rôle de service statistique ministériel chargé d'éclairer les prises de décision de politique éducative. Nous verrons dans un troisième temps que la multiplicité des évaluations menées par la DEPP, parmi lesquelles PISA éclairé par les travaux nouveaux menés dans le cadre de CEDRE, a également été valorisé auprès du monde universitaire dans une volonté de conférer aux travaux de la DEPP une dimension proche des démarches de recherche, afin de renforcer leur légitimité d'expert indépendant de l'évaluation en éducation.

1) L'invisibilisation de l'expertise de la DEPP à l'origine de critiques de fond de PISA

« Aucune conséquence politique quand PISA a commencé, au début des années 2000. La France a cassé le thermomètre, a critiqué l'instrument de mesure, et n'en a tiré aucune conséquence décisionnelle. »³³⁰

Par ces mots, Catherine Moisan, directrice de la DEPP au début des années 2010 et conseillère technique de plusieurs ministres de l'éducation nationale dans les années 1980 et 1990, résumait lors de notre entretien sa perception de l'utilisation des premiers résultats de PISA par les acteurs français des politiques éducatives, y compris les agents de la DEPP. Son regard sur la réception de PISA par la DEPP et les ministres de l'Education nationale entre en résonance avec la façon dont l'usage des premiers cycles de PISA, en termes de pilotage de la politique éducative française, paraît rétrospectivement faible aux yeux des observateurs, chercheurs, journalistes, qu'ils soient nationaux ou internationaux. Malgré le travail fourni par les équipes du ministère en amont de son lancement en 2000, et l'implication de la direction dans la conduite de l'évaluation par son National Project Manager et des experts sollicités pour travailler sur la confection des items de test, les résultats de la France à PISA n'auraient pas été suivis de réflexions sur ce qu'ils pouvaient montrer du fonctionnement du système éducatif français et impliquer comme réformes ou ajustement de ce fonctionnement. Ou du moins, ils ne l'auraient pas été avec l'intensité nécessaire. En outre, au-delà de ce non-usage, les évaluateurs français auraient largement critiqué la validité de la méthode d'évaluation à laquelle ils avaient pourtant participé, au même titre que l'ensemble des autres pays participants et issue d'un consensus obtenu après plusieurs années de travail. Néanmoins, et nous le verrons dans la suite de ce chapitre, ce discours unifié des évaluateurs nationaux, se double en arrière-plan d'un travail de construction d'une nouvelle évaluation qui reprend en fait certains des principes de PISA pourtant sujets à ces fortes critiques. Pour mieux saisir la stratégie des évaluateurs nationaux dans ces réceptions parallèles et à

³³⁰ Entretien avec Catherine Moisan, Directrice de la DEPP de 2013 à 2015, 10 juin 2020.

première vue difficilement compatibles, nous observerons d'abord quelles ont été les critiques émises par la DEPP et dans quelle mesure elles font écho à leur implication manquée évoquée au chapitre précédent.

Lors du lancement de PISA, plusieurs choix méthodologiques font l'objet de lourdes critiques au sein de la DEPP. Deux de ces critiques vont structurer le discours des agents de la DEPP. La première porte d'abord sur le fondement même de PISA : le recours à la psychométrie. Elle se cristallise sur l'utilisation de questions à choix multiples. La seconde concerne une caractéristique intrinsèque de l'évaluation internationale : la traduction des items en plusieurs langues, qui revêt une importance particulière compte-tenu de l'évaluation nouvelle en France du niveau de *littératie* des élèves. Ces questions méthodologiques vont susciter une attention particulière de la part de la DEPP dès les premiers travaux de construction de PISA. Ses agents tentent ainsi de faire valoir leur expertise en discutant des points de méthode, en sollicitant des experts pouvant concourir à la création des items ou à leur traduction, ou encore en alertant sur ce qui constituait, selon eux et leur expérience de l'évaluation en éducation, des erreurs méthodologiques lourdes limitant la validité des résultats obtenus.

Ces interventions françaises n'ont pas abouti à des changements méthodologiques de la part de l'équipe PISA-OCDE. A la sortie des premiers – et moins bons qu'espérés – résultats des élèves français, ce sont les points de méthode peu maîtrisés par les évaluateurs de la DEPP qui seront la cible de leurs critiques et de la remise en question des conclusions de l'OCDE. Nous entendons ici l'absence de maîtrise sous deux sens complémentaires. D'une part, il s'agit de ce qui peut s'apparenter à un défaut, ou à un « trou », dans l'expertise des agents de la DEPP, parce que PISA utilise des méthodes que la DEPP n'emploie pas dans ses évaluations nationales, pour des raisons de traditions d'évaluation divergentes, et qu'elle connaît de fait peu aussi bien théoriquement que de façon pratique. D'autre part, il s'agit de ce sur quoi la DEPP n'a pas eu de prise : les processus décisionnels menant aux choix méthodologiques de PISA dans lesquels la DEPP n'a pas réussi à imposer son point-de-vue, son œil expert.

a) Une critique politique de la méthode sous couvert d'une bataille de modèle d'évaluation

Avant de développer spécifiquement les grands points de méthode de PISA cibles de la critique de la DEPP, il faut les replacer dans le cadre plus large de l'opposition de deux modèles d'évaluation dans laquelle ils s'inscrivent.

Les critiques méthodologiques énoncées par la DEPP se font avec en toile de fond une critique politique du modèle d'évaluation choisi pour PISA. PISA s'est construit sans l'aval inconditionnel de la DEPP, mais dans les règles de fonctionnement des instances décisionnelles de PISA fondées sur le principe de consensus de l'ensemble des pays participants. Ce principe de consensus entre plusieurs dizaines de pays participant à PISA, qui implique l'acceptation d'une méthode d'évaluation commune à tous et différente de celle voulue par chacun, est source d'enjeux forts dans le cadre d'évaluations des systèmes éducatifs. Le consensus doit en effet néanmoins garder l'exigence de représentativité nécessaire à la validité de l'évaluation. C'est précisément dans le processus de construction du consensus sur les principes de l'évaluation que naît l'opposition méthodologique principale exprimée par la DEPP : celle d'une absence de prise en compte par le modèle psychométrique de PISA des particularités de l'évaluation du système éducatif français dont elle a une longue expérience. L'imposition ressentie par les agents de la méthode psychométrique et du processus complexe de passation des tests et du traitement des données, est perçue en interne comme une manifestation d'un pouvoir politique de la part de l'organisation internationale. L'OCDE, et en arrière-plan son idéologie de l'évaluation proche du néo-libéralisme soutenu par les pays anglo-saxons et les Etats-Unis en particulier, étendrait avec PISA le modèle d'évaluation américain à l'ensemble des pays participants, aux dépens des traditions d'évaluation nationales et française en particulier.

Une note interne, anonymisée, transmise par l'une des personnes que nous avons rencontrées en entretien, résume en deux pages la stratégie de remise en question des principes de fond de l'évaluation qui a primé au début de

l'évaluation. Intitulée « Critique et enjeux des enquêtes internationales sur les acquis des élèves : à propos de l'enquête PISA », elle a été rédigée par un agent de la DEPP et adressée au directeur de la DEPP à sa demande, après la deuxième édition de PISA. En quelques paragraphes tous les éléments de critique de PISA y figurent, sans que l'évaluation ne soit clairement nommée, si ce n'est dans le titre général du document. L'évaluation conduite par l'OCDE est plutôt englobée dans un ensemble plus large « *d'enquêtes internationales* » qui présentent pour son auteur des problématiques à la fois technique, culturelle et politique.

Cette perception du poids politique de PISA à travers son modèle évaluatif transparait dès l'introduction de la note. Elle commence par énoncer les difficultés inhérentes à la critique des évaluations internationales : « *elle est délicate d'une part car leur utilité n'est pas contestable – il faut les faire – et d'autre part car leurs méthodologies sont complexes de ce fait, la critique ne peut se faire qu'à un niveau souvent très technique* ». La formule « *il faut les faire* » témoigne bien de la tension autour de la participation aux évaluations internationales et, dans ce cas, à PISA. Si la France participe, c'est qu'elle n'a pas le choix. L'évaluation est imposée par l'organisation internationale qui use du pouvoir d'influence dont elle dispose, et subie par les évaluateurs de la DEPP qui ne peuvent qu'y participer en faisant valoir au mieux leur expertise en matière d'évaluation.

« PISA de toute façon c'est un mal nécessaire. Il faut participer. On ne pouvait pas ne pas le faire, alors que la France est la cinquième puissance mondiale et que l'OCDE est chez nous. On est obligés, alors on a fait ce qu'on a pu pour être écoutés et ne pas laisser le modèle anglo-saxon prendre le dessus. »³³¹

« L'OCDE est en France, c'est une grosse pression. On est un des gros pays de l'OCDE, alors une des cinq premières puissances qui ne veut pas participer à un projet ça interroge, c'est tout de suite compliqué pour recruter les autres. »³³²

³³¹ Entretien avec Anne-Laure Monnier, Chargée de mission au Conseil supérieur des programmes, 12 septembre 2016.

³³² Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 18 mars 2019.

La formule de « *mal nécessaire* » est représentative de la façon dont l'arrivée de PISA dans le panel d'évaluations de la DEPP a été accueillie au départ par les agents. En interne, la présence du siège de l'OCDE à Paris est vue comme un facteur de pression supplémentaire, qui s'ajoute à la pression forte non seulement de participer activement aux modules d'évaluation proposés par l'OCDE mais aussi de se présenter comme l'un des moteurs de l'organisation internationale, du fait du poids politique de la France allant de pair avec la contribution financière de base importante de la France, proportionnelle à son PIB.

Le fonctionnement jugé descendant des réunions où sont prises les décisions stratégiques autour de PISA suscite une réaction critique de la part des évaluateurs français.

« Moi j'ai travaillé sur la deuxième vague de PISA. On a fait au moins trois meetings internationaux. Il faut arriver à convaincre les gens qui y sont, parce que même si PISA c'est obligatoire il y a quand même des difficultés techniques qu'on voit bien. Et ça c'est vraiment compliqué dans les réunions. »³³³

« Au début de PISA, on a eu des problèmes techniques, c'est normal. On propose des solutions à ces problèmes, de faire par exemple une autre enquête pour tester des choses. Nous on est très favorables à ça, de toute façon c'était un début, tout était à créer, il y avait plein de choses à décider, les pays cherchaient un consensus. C'est très attirant tout ça. Mais bon, au final tout est sous l'égide de l'OCDE. »³³⁴

Alors même qu'ils disposent d'idées à expérimenter et de modèles solides et éprouvés au cours de leur longue pratique de l'évaluation des résultats des élèves, il leur est difficile d'influer sur les décisions prises lors des réunions. L'impression partagée est celle d'une part d'une orientation forte des débats par les équipes en charge de PISA pour l'OCDE et d'autre part de choix finaux décidés d'avance,

³³³ Entretien avec Anne-Laure Monnier, Chargée de mission au Conseil supérieur des programmes, 12 septembre 2016.

³³⁴ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 21 mars 2019.

sur lequel il est presque impossible d'influer malgré toute l'argumentation possible.

Ces critiques fortes sur l'absence de « contre-pouvoir » des pays participants à PISA sont peu visibles dans les interventions officielles des représentants de la France auprès de l'équipe PISA-OCDE lors des réunions synthétisées par écrit et disponibles dans les archives de l'organisation. Si l'on regarde les comptes-rendus des réunions du GPP, ils ne mentionnent que très rarement les oppositions exprimées par les présents mais rendent davantage compte des consensus obtenus en fin de réunion. Certes, l'absence de mention d'opposition dans un compte-rendu de réunion ne signifie pas pour autant absence d'opposition signifiée en réunion. Mais il est difficile de savoir dans quelle mesure la synthèse exigée par l'organisation internationale pour les comptes-rendus a pu faire disparaître les oppositions et interrogations formulées par les représentants des pays, dont la France, et dans quelle mesure le jeu d'influence a pu se jouer majoritairement en dehors des temps de réunion à proprement parler comme cela a pu nous être rapporté dans certains de nos entretiens.

En tout état de cause, si aucun de nos interlocuteurs n'a pu nous indiquer si les représentants français dans les différentes instances décisionnaires de PISA avaient bel et bien exprimé publiquement, lors des réunions et des groupes de travail, leur remise en question de certains principes, plusieurs ont bien relaté la façon dont les échanges entre les pays participants dépassent les frontières physiques et temporelles des temps de travail formels prévues par l'OCDE. Les NPM et représentants français au Groupe des Pays Participants échangent régulièrement entre eux tout au long de l'année, et cela depuis le début de PISA, ce qui permet de faire part d'éventuels doutes sur des choix méthodologiques à prendre lors de prochaines réunions organisées par l'OCDE. Ces temps entre deux réunions ont permis à plusieurs reprises la construction d'une voix commune et de jeux d'alliances pour faire part auprès de l'équipe de l'OCDE, lors des réunions du GPP, des réticences à voir appliqués tels ou tels principes méthodologiques. L'un des exemples les plus parlants est celui du recueil de la profession des parents des élèves. Cette information est cruciale dans le

déroulement et le traitement des réponses puisqu'elle permet de construire l'indice de situation socio-économique et est utilisée notamment pour rendre compte de l'effet de la scolarité sur les trajectoires des élèves par rapport à leurs parents, et donc la reproduction, la réduction ou le renforcement des inégalités sociales par l'école. La méthode validée par l'équipe PISA-OCDE se fondait sur une auto-déclaration par les élèves de la profession de leurs parents, ce à quoi les agents de la DEPP présents aux réunions techniques se sont tout de suite opposés. Faute d'obtenir l'écoute de leurs interlocuteurs de l'OCDE après une première alerte, ils ont cherché à coordonner les oppositions qu'ils avaient pu entendre chez leurs homologues d'autres pays participants.

« On s'était dit on va demander aux élèves la profession des parents pour entrer dans PISA. J'avais lu un rapport de l'EHESS qui disait que les élèves n'étaient pas très au courant de la profession réelle de leurs parents. On n'en discute jamais dans les familles. Ils ne savaient pas vraiment, surtout pour la formation, 30 ou 40% ne savaient pas. Avec Thierry Rocher, on a dit « non ça ne marche pas ». En France elle ne marche pas en tout cas. L'OCDE a dit « mais si, c'est pareil partout », j'ai dit qu'il y avait une étude de l'EHESS. On a été coordonnés. Le Japon a dit comme nous, du coup bon. La Tchécoslovaquie a dit ça aussi, l'Allemagne a dit ça aussi. Et là l'OCDE a dit que c'était déjà parti comme ça. Je suis monté au créneau, je me suis un peu énervé. J'ai demandé à ce qu'on fasse une expérimentation, et ils ont dit oui. On a collecté la profession des parents, comme demandé par l'OCDE. On a demandé à l'INSEE d'appeler les parents pour comparer. En vrai c'était pas utilisable ce que disaient les élèves. L'OCDE a dit « on ne va pas changer, c'est prévu comme ça », ils ont dit « on va changer de grille ».

Dans le consortium il y avait quelqu'un qui travaillait pour le MEN aux Pays-Bas. Ils ont dit « on va fournir un expert des Pays-Bas, qui a trouvé une échelle, au lieu des catégories sociales ». Alors que nous on proposait quelque chose ! On disait de demander aux chefs d'établissements. Ça sera beaucoup plus sûr que de demander aux élèves. Mais nan ça ne rentrait pas dans le projet, peut-être pas dans les sous. Il fallait que ça soit les élèves qui donnent ça. Peut-être parce qu'ils collectaient les infos eux-mêmes.

C'est de la folie pure. Un boulanger c'est mieux qu'un boucher ? Il est venu exposer ça le chercheur néerlandais. Aux Pays-Bas, peut être que c'est mieux, mais au Japon ? Alors que ça devait marcher partout. Mais quand on a une échelle continue, les différences sont moins grandes que quand on

fonctionne par catégorie. En 2000 ça a été appliqué comme ça, avec des tas de problèmes.

Ce qui n'est pas normal c'est qu'on ne dise pas qu'on a un problème, et qu'on continue. Ils font ça parce que c'est pas vendeur. »³³⁵

Deux dimensions sont intéressantes à relever dans la façon dont est narré cet exemple et illustrent la perception d'une absence de possibilité d'apport d'expertise par la DEPP à l'OCDE. D'abord, l'expertise individuelle des agents de la DEPP impliqués dans PISA, aux compétences actualisées sur les méthodes employées dans le cadre des évaluations des résultats des élèves, est ici décrite comme centrale par l'un des protagonistes au cœur du conflit entre la DEPP et l'OCDE au sujet du recueil de la profession des parents des élèves participant à PISA. Les évaluateurs de la DEPP ont pu repérer les potentielles failles de la méthode proposée, demander une expérimentation préalable au niveau national pour voir si leurs hypothèses étaient valides et leurs oppositions légitimes. Malgré le succès de l'expérimentation, la méthode n'a pas réellement été revue, mais seulement modifiée à la marge par l'équipe PISA-OCDE. En impliquant un pays participant dans l'ajustement de la méthode, l'OCDE a conservé le fonctionnement auquel elle tient de la contribution des pays dans la conception de PISA. Par cela, elle a rendu difficile l'argument de l'imposition par l'organisation de ses choix de méthode, puisque la méthode finale a été proposée par un pays et validée par les autres.

La deuxième dimension intéressante est celle attribuée à l'OCDE, celle d'une organisation dont la légitimité tient à la justesse de ses propositions « du premier coup ». Pour notre évaluateur, la reconnaissance par l'OCDE de l'inadéquation d'une de ses propositions vaut perte de légitimité. La lecture de l'imposition du modèle et de la méthode d'évaluation voulus par l'OCDE est bien celle d'un conflit de légitimité entre l'organisation internationale et les évaluateurs nationaux des pays participants.

³³⁵ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 18 mars 2019.

L'imposition du modèle psychométrique est d'autant plus source de craintes pour la DEPP qu'il est associé au modèle d'évaluation anglo-saxon et à l'introduction en particulier du secteur privé dans l'évaluation des résultats de l'éducation nationale, vue comme une perte forte d'indépendance des évaluateurs et un abandon de l'intérêt général comme moteur de l'évaluation. La crainte est exprimée notamment dans la note anonymisée de la DEPP. L'auteur voit, avec le fonctionnement acté de PISA, les opérateurs privés de l'évaluation en éducation prendre la main sur les modèles d'évaluation appliqués par les évaluateurs publics, dépourvus de tout intérêt financier ou commercial par la nature de leur statut.

« La prise en main par l'OCDE d'une partie de ces enquêtes et l'arrivée d'institutions nouvelles a mis en évidence les enjeux politiques et économiques de la maîtrise d'œuvre de ces enquêtes. Il s'agit du marché de l'évaluation et en arrière-plan celui de l'éducation, puisque l'évaluation permet de juger de la qualité des formations dispensées voire de les certifier. A titre d'exemple citons les faits suivants :

. un organisme comme ETS (Educational Testing Service), organisme privé américain, qui s'est construit un empire politique et commercial sur les tests d'admission dans les universités américaines et auteur du TOEFL (Test of English as Foreign Language) cherche maintenant à développer une politique commerciale internationale (ETS a ouvert une représentation en Europe).

. on assiste à un phénomène de monopole et de renvoi d'ascenseur entre institutions et personnes. Les institutions sont présentes dans les différents organismes : ainsi la Commission Internationale des Tests (ONG) a élaboré des recommandations sur l'adaptation et la traduction des tests. Le colloque au cours duquel ont été présentées ces recommandations était animé par la Commission Internationale des Tests mais aussi par ETS avec la participation de Microsoft. A ce colloque intervenaient des universitaires à la fois membres de la commission des tests et participants à PISA. Les recommandations garantissent que les produits satisfont à des normes de qualité, on voit donc le danger de la collusion entre une organisation professionnelle garante de la qualité des instruments d'évaluation et un éditeur de test.

Il est important de dire que ces enquêtes internationales, sont une vitrine pour un certain nombre d'organisations afin d'apparaître comme les meilleures, voire les seuls, dans ce domaine. Les mêmes institutions ont fait plusieurs tentatives pour se placer auprès de l'UNESCO. On peut, il me

semble, faire l'hypothèse – sans que cela apparaisse pour de la paranoïa antimondialisation et anti-américaine – qu'il est en train de se développer un impérialisme culturel dans ce domaine de l'évaluation. Signalons qu'il existe déjà largement dans le domaine de l'évaluation psychologique puisque – pour parler rapidement – on mesure « l'intelligence » des petits français avec des adaptations de tests américains et la personnalité des cadres français avec des questionnaires adaptés de questionnaires américains. »

Ces inquiétudes exprimées en interne à la DEPP témoignent de la méfiance envers ce qu'ils identifient comme une marchandisation des évaluations des résultats de l'éducation, et qu'ils ne souhaitent pas voir supplanter leurs propres travaux réalisés en toute indépendance d'intérêts politiques ou d'intérêts privés.

Ces craintes se renforcent, du moins au niveau des évaluateurs français chargés de s'assurer de l'excellence technique de PISA, après la fin des années 2000, au moment du morcellement de la gestion des tests et de la multiplication des interlocuteurs du consortium chargé de prendre en charge la construction, la coordination internationale de la passation des tests et le traitement des données recueillies. L'arrivée de nouveaux acteurs étiquetés comme faisant partie du secteur privé a pu poser question sur leurs motivations. Cela a renforcé l'attention portée par la DEPP à la conception technique des tests, alors même que leurs marges de manœuvre se trouvaient de plus en plus réduites. L'un des agents de la DEPP impliqués dans la conception des évaluations de PISA sur les mathématiques a exprimé en entretien sa perception de l'évolution parallèle du type d'acteur du consortium et des capacités d'intervention des NPM auprès de l'OCDE.

« Depuis 2015 c'est plus morcelé. Le cadre maintenant c'est Pearson qui a tout pris en charge. Pearson qui sont des vendeurs de livres, faut pas oublier ça. Tandis qu'avant c'étaient des entreprises qui étaient semi-publiques. ACER partiellement, ce n'est pas qu'une entreprise privée, elle travaille pour les états. Et donc à partir de 2015, il y a eu une manière de travailler très différente, impliquant moins les NPM dans la prise de décision. Les meetings étaient plus des moments d'information que des moments de partage de pratiques ou d'idées. D'ailleurs ils étaient plus courts. Et on a pu ressentir une sorte de fronde des NPM, une fronde qui n'a pas vraiment

porté ses fruits dans la mesure où on n'a pas vraiment prolongé au niveau de PISA governing board. »³³⁶

On perçoit clairement dans cette déclaration toute la dichotomie entre évaluation publique et évaluation privée, et la nécessité dans le cas du modèle français de l'évaluation de l'éducation d'assurer qu'il ne se double pas d'un intérêt économique qui serait incompatible avec sa mission de service public. Avec des acteurs comme Pearson, ayant pour objectif de « vendre des livres », l'évaluation des résultats de l'éducation par PISA n'a plus pour objectif unique de fournir aux pays les données les plus fiables pour orienter l'action publique en éducation, mais pourrait devenir un marché à se partager pour des acteurs privés devant assurer une rentabilité maximum. Alors que la DEPP cherche à rassurer au niveau national sur l'indépendance de ses évaluations du pouvoir politique, elle cherche également à alerter sur les risques de voir les évaluations à laquelle elle participe se lier à des intérêts financiers et affirmer son indépendance du secteur privé.

Ces oppositions de modèles et les craintes de voir les évaluations en éducation obéir davantage à des impératifs de rentabilité économique voire des velléités d'hégémonie que de servir à l'amélioration des systèmes éducatifs, renforcent l'attention portée dès le lancement de PISA à tous ces détails techniques. La bataille de modèle vient s'ajouter aux souvenirs d'une expérience malheureuse pour la France dans ses relations avec l'OCDE, immédiatement antérieure au lancement de PISA. Elle vient renforcer la méfiance des évaluateurs nationaux concernant la fiabilité des résultats à venir et les incite à regarder avec une attention critique la façon dont l'évaluation est construite.

³³⁶ Entretien avec Franck Salles, Chargé d'étude sur l'évaluation des élèves, National Project Manager, DEPP, 11 juillet 2017.

b) Le traumatisme de IALS : un précédent qui vient renforcer les critiques de la méthode de PISA

L'expérience malheureuse de l'enquête IALS, *International Adult Literacy Survey*, lancée par l'OCDE en 1994, a eu un effet sur la façon dont les évaluateurs de la DEPP ont eu de réagir aux problèmes méthodologiques posés selon eux par PISA. A l'époque, l'OCDE avait souhaité lancer une évaluation des niveaux de littératie des adultes de plus d'une vingtaine de ses pays membres, en utilisant comme matériau support des documents de la vie courante. L'organisation internationale s'était alors appuyée sur des entreprises anglo-saxonnes pour mener l'évaluation, et en particulier ETS que l'on retrouvera ensuite dans le consortium chargé de PISA. Les résultats français avaient été particulièrement mauvais, puisque 40% de la population adulte était considérée à la fin de l'évaluation comme non seulement incapable de déchiffrer un texte court, mais aussi d'en saisir le sens. Ce chiffre était en outre bien plus élevé que ceux de nos voisins européens. Or, de nombreux problèmes méthodologiques se posaient et l'évaluation avait été considérée comme peu crédible³³⁷. Le ministère de l'Éducation nationale s'était opposé à la publication des chiffres français et les médias s'étaient rapidement emparés de cette polémique pour la placer au devant du débat public, aboutissant au retrait de la France de l'évaluation de façon définitive.

Au sein de la DEPP, plusieurs des agents engagés dans l'évaluation PISA que nous avons rencontrés au milieu des années 2010 gardaient un souvenir vif de l'« histoire » de IALS, à laquelle ils ont spontanément pensé en évoquant PISA.

« Cette histoire de IALS, ça a été désastreux pour la France. La France s'est retirée trop tard. La méthodologie était très mauvaise, la France se retrouvait avec 40% d'illettrés. Du coup après Thélot a demandé des expertises. »³³⁸

³³⁷ Louis-André Vallet, « L'enquête Information et vie quotidienne 2011 parmi les dispositifs nationaux et internationaux de mesure des compétences des adultes », *Economie et statistique*, n°490, 2016.

³³⁸ Entretien avec Bruno Trosseille, Directeur du bureau des évaluations des élèves, DEPP, 5 janvier 2016.

« La France a été très critique sur l'échantillonnage de IALS et la façon dont l'enquête était menée dans les foyers, et spécialement la motivation des individus pour faire le travail de lecture car ils n'étaient pas captifs. Les enquêteurs quittaient le domicile au lieu de rester pendant le remplissage des questionnaires. Et pourtant on l'a dit ! »³³⁹

La méthode de passation avait été jugée à l'époque trop peu rigoureuse pour pouvoir fournir des résultats représentatifs du niveau des adultes participant à l'évaluation et en tirer des conclusions sur le niveau général de la population adulte française. Alors que la DEPP, lors de ses enquêtes au domicile des personnes, organisait la passation des tests en présence de l'enquêteur, IALS permettait à l'enquêté de rester seul devant le test. Cette différence avait été pointée du doigt par la DEPP au moment de la mise en place de la méthode de IALS sans que cela ne fasse l'objet d'une révision méthodologique pour le choix de la technique de passation finale, ni que cela inquiète suffisamment le Ministère pour qu'il se retire dès le départ de l'évaluation internationale. L'absence de prise en compte de cette alerte pourtant formulée par la DEPP, qui avait contribué à aboutir à cette situation de crise au moment de la sortie prévue des résultats de IALS, est représentative de l'influence réduite que pouvait avoir la DEPP à l'époque, à la fois auprès de l'OCDE et des décideurs politiques.

Une fois la France retirée de l'évaluation à l'initiative du ministre de l'époque, François Bayrou, des études ont été menées à la demande de Claude Thélot, alors directeur de la DEPP, pour mieux saisir les points d'achoppement méthodologiques à retravailler pour les évaluations à venir. Ces travaux menés en interne ont notamment servi à construire l'enquête de l'INSEE *Information et Vie Quotidienne* conduite à partir de 2004 en collaboration notamment avec la DEPP, l'Agence Nationale de Lutte Contre l'Illettrisme³⁴⁰. Un important débat

³³⁹ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 18 mars 2019.

³⁴⁰ L'ensemble des contributeurs à l'enquête INSEE IVQ sont l'INSEE, l'Agence nationale de lutte contre l'illettrisme, la DEPP, la Dares, le Département des statistiques, des études et de la documentation du ministère de l'Intérieur, la Délégation générale à la langue française et aux langues de France du ministère de la Culture, et le Secrétariat général du comité interministériel des villes.

a également émergé au sein du monde universitaire, en lien avec les travaux demandés par Claude Thélot, sur les processus de validation d'évaluations internationales de ce type, en les replaçant dans les oppositions françaises historiques aux méthodes psychométriques déjà évoquées³⁴¹. Deux paramètres techniques centraux ont été identifiés comme ayant eu un effet de biais sur le très mauvais résultat français, le rendant non significatif : le mode de passation qui aurait engendré une moindre motivation des enquêtés, ou tout du moins une motivation inégale ; et la traduction des questions dont la difficulté en langue française aurait été supérieure à celle d'autres langues, et notamment la version initiale en anglais.

« L'histoire » ou l' « affaire IALS ³⁴² » comme ont pu la nommer les chercheurs par la suite, a engendré plusieurs mécanismes de justification de la part des évaluateurs de la DEPP face à toute évaluation aux résultats jugés potentiellement peu valorisants pour le système éducatif français, et en particulier lorsqu'elle est menée par la même organisation, l'OCDE. Le premier mécanisme est celui d'une remise en question, pour ne pas dire mise en cause, de la méthode choisie pour les évaluations de l'OCDE qui aurait été conservée, alors que les évaluateurs de la DEPP auraient averti de sa faiblesse. En plus du problème de passation, que la DEPP avait relevé, deux autres éléments de IALS repris dans PISA ont ainsi suscité la méfiance avant même la sortie des résultats : le besoin de traduire à nouveau les items d'évaluation de compétence en littérature, avec le risque que les difficultés de compréhension d'une langue à l'autre soient inégales ; et le recours à des prestataires anglo-saxons éloignés des espaces d'influence des évaluateurs français et sur lesquels l'expérience avait montré qu'il était complexe de faire pression pour que les alertes soient prises en compte.

Le second mécanisme est celui de la réaffirmation du rôle d'expertise méthodologique de la DEPP auprès des décideurs politiques français

³⁴¹ Voir l'ouvrage de référence sur la crise qui a entouré l'évaluation IALS en France, Alain Blum et France Guérin-Pace, *Des lettres et des chiffres : des tests d'intelligence à l'évaluation du « savoir lire », un siècle de polémiques*, Paris, Fayard, 2000, 191pp. Pour un rappel du parcours manqué en France de la psychométrie, voir chapitre 1 de cette thèse.

³⁴² Alain Blum et France Guérin-Pace, *op. cit.*

notamment. Ce qui est vécu par les agents de la DEPP en poste à l'époque comme une véritable « crise IALS » est aussi dans les discours de ces mêmes agents un exemple de la façon dont la DEPP a pris en main, pour le ministère, l'imposition de leur expertise en matière de fiabilité des évaluations face à l'OCDE et aux autres participants. Jean-Claude Emin, ancien sous-directeur à l'évaluation au sein de la DEPP accorde à la DEPP cette capacité d'apprentissage et de questionnement qui aurait fait d'elle un acteur évaluateur aux productions particulièrement robustes.

« La DEP a progressivement développé de nombreux travaux et études, concentré en son sein la mise au point et la diffusion de la plus grande partie des instruments de la politique d'évaluation éducative et organisé l'animation, voire le pilotage, de nombreux travaux de recherche en la matière. (...) Citons (...) qu'elle a pris en main la participation française aux évaluations internationales et a joué un rôle important dans l'évolution de ces évaluations, y compris lors de crises, telle celle de l'enquête IALS sur les compétences des adultes, crise à l'occasion de laquelle ont été posées par la France de nombreuses questions touchant à la fiabilité des évaluations internationales en éducation, ce qui a permis d'améliorer sensiblement la technologie – et donc la qualité – de ces évaluations. »³⁴³

Pour Jean-Claude Emin, la crise liée à IALS a permis aux évaluateurs nationaux de prendre l'initiative de se saisir pleinement et en toute autonomie, sans commande ministérielle, de sa mission d'expertise internationale en statistiques et indicateurs sur l'éducation³⁴⁴ que lui confère son statut de service statistique ministériel. La lecture que fait Jean-Claude Emin de « l'affaire IALS » avec le recul des années est celle d'une preuve de maîtrise technique des méthodes d'évaluation, y compris lorsqu'elles impliquaient des outils d'évaluation peu pratiqués en France comme la littératie ou la construction d'items comparables d'une langue à l'autre en terme de difficulté de réponse. Ce travail d'alerte sur les manquements méthodologiques aux conséquences graves sur la fiabilité de

³⁴³ Jean-Claude Emin, « Réflexions sur le rôle de la direction de l'évaluation et de la prospective (DEP) en France », *Education et Sociétés*, 2012/2, n°30, p.51.

³⁴⁴ Définition des missions de la DEPP en tant que service statistique ministériel fournie par l'INSEE, consultable au lien suivant : <https://www.insee.fr/fr/information/2113161>

l'étude a selon lui posé les fondements de réflexions plus larges sur les évaluations futures de l'OCDE en éducation, et en particulier PISA.

c) La critique des outils psychométriques

Alors que le contexte du retrait de la France de IALS a renforcé la méfiance des évaluateurs de la DEPP envers à la fois la méthode de non-concertation de l'OCDE et la représentativité des évaluations qu'elle menait sur les résultats de l'éducation, un point de méthode a particulièrement cristallisé les oppositions lors du lancement de PISA. La bataille de modèles d'évaluation s'est focalisée largement autour de l'utilisation des méthodes psychométriques, considérée comme la méthode anglo-saxonne historique pour la réalisation de leurs évaluations nationales, et en particulier celles des Etats-Unis.

Du point-de-vue technique, les outils psychométriques supposent l'utilisation du questionnaire à choix multiple (QCM) qui n'était pas à l'époque utilisé dans les évaluations françaises, plutôt enclines à recourir à des questions laissant aux élèves la possibilité d'écrire³⁴⁵. Le format du QCM fait d'ailleurs régulièrement l'objet de travaux de recherche didactiques ou techniques, pour interroger la qualité de ce format de question en proposant des aides logistiques pour réaliser des analyses statistiques pertinentes des résultats obtenus³⁴⁶, construire des questions permettant d'évaluer ce que l'on souhaite et obtenir des scores significatifs³⁴⁷. Le format de la question n'implique pas uniquement de déterminer le contenu de ce sur quoi on souhaite interroger les élèves, mais il

³⁴⁵ Voir par exemple, Nadine Grapin et Nathalie Sayac, « Evaluer la maîtrise de la numération écrite chiffrée : choix du QCM et validité d'items d'évaluations externes », *Education et didactique*, 11-3, 2017.

³⁴⁶ Jean-Marc Braibant, « Comment aider et accompagner les enseignants universitaires à améliorer la qualité des examens par QCM sur base des analyses d'items ? » *Actes du 27^e Colloque de l'ADMEE Europe*, 2015, Liège.

³⁴⁷ Voir les travaux de D. Leclercq, *La conception des questions à choix multiples*, 1986, Bruxelles, Labor ; *Qualité des questions et signification des scores*, 1987, Bruxelles, Labor ; Nathalie Sayac et Nadine Grapin, « Evaluer les capacités des élèves à résoudre des problèmes dans le cas d'une évaluation externe en France : les spécificités de la forme QCM », *Education et francophonie*, XLII(2), 64-83, Nathalie Sayac et Nadine Grapin, « Evaluer par QCM en fin d'école : stratégies et degrés de certitude », *Annales de didactique et de sciences cognitives*, 19, 2014.

nécessite aussi de mobiliser certaines compétences pour se représenter le problème posé³⁴⁸, vérifier et justifier une réponse³⁴⁹, pour réduire autant que possible les réponses au hasard, prendre en compte la probabilité de répondre correctement au hasard plutôt que de risquer une mauvaise réponse lorsqu'on n'est pas sûr de son raisonnement, etc. Au lancement de PISA, la méthode était jugée, par les agents de la DEPP impliqués dans PISA, comme trop méconnue par les élèves français, comparés à leurs homologues américains, canadiens, australiens ou britanniques. Une crainte avait été exprimée dès le départ sur le biais que cette absence d'habitude du QCM chez les élèves français pouvait susciter et le souhait avait été exprimé d'en tenir compte dans l'analyse des scores au moment de construire les groupes de compétence à partir des résultats de tous les pays. Finalement, le biais avancé n'a pas suscité de travail spécifique de la part de l'OCDE ou du consortium chargé de produire les scores pour l'analyse. Cette absence de réaction de la part de l'organisation internationale n'a pas entraîné davantage de déclarations de la part de la DEPP. En effet, la critique de l'inadaptation du format de QCM a en réalité rapidement été neutralisée dans les faits car les analyses des réponses des élèves français ont montré dès le départ qu'ils répondaient mieux aux questions fermées de type QCM qu'aux questions ouvertes. Nous verrons même par la suite que cette critique du QCM utilisé dans les évaluations du système éducatif français a été menée presque simultanément à une démarche d'intégration des méthodes psychométriques dans les évaluations nationales de la DEPP et notamment de nouvelles créations très inspirées de PISA, comme l'évaluation nationale CEDRE.

Du point-de-vue culturel et politique, les modèles psychométriques de PISA s'inscrivent dans la bataille de modèles d'évaluation entre la DEPP et l'OCDE déjà évoquée précédemment, et matérialisent la crainte d'une grande proximité entre l'évaluation PISA et le « modèle américain ». Ce modèle serait caractérisé

³⁴⁸ Jean Julo, *Représentation des problèmes et réussite en mathématiques*, 1995, Rennes, Presses universitaires de Rennes, et « Des apprentissages spécifiques pour la résolution de problèmes ? », *Grand N*, 69, 2002.

³⁴⁹ Sylvie Coppé et Catherine Houdement, « Réflexion sur les activités concernant la résolution de problème à l'école primaire ». *Grand N*, 69, 2002.

bien sûr par une tradition d'évaluation très différente de la tradition française, mais aussi comme on l'a vu, par une prédominance du secteur privé à but lucratif et une volonté d'homogénéisation des pratiques – américaines bien sûr –, et en particulier des méthodes psychométriques. La note anonymisée de la DEPP énonçait ainsi à ce sujet :

« Pour des raisons institutionnelles, les méthodes adoptées sont principalement celles qui ont été développées aux Etats-Unis pour des comparaisons fédérales. Les organismes travaillant à ces évaluations ont cherché à étendre leur champ d'action aux comparaisons internationales. On peut se demander dans quelle mesure ce cadre peut être valablement étendu aux comparaisons internationales. Par voie de conséquence, les travaux d'autres pays, d'autres organismes, d'autres chercheurs ont été soit laissés de côté soit minimisés. Il faut, contrairement au discours que tiennent ces organisations, se garder de considérer que la méthode adoptée est la seule et la meilleure.(...)

Bien que chaque pays puisse proposer des exercices, la majorité des épreuves est cependant d'origine anglo-saxonne (voire reprises des épreuves d'évaluation des jeunes américains) ou sinon calquées sur des modèles anglo-saxons ».

La sémantique est bien celle de la bataille entre plusieurs modèles dans laquelle la France se serait engagée avec peu de succès et qui aurait été remportée par les Etats-Unis. Plus gros contributeurs financiers de l'OCDE puisque les contributions sont calculées en pourcentage de PIB, perçus comme particulièrement influents sur les orientations libérales prises par l'Organisation internationale dont ils sont à l'origine de la création, ils sont considérés, dans cette note au directeur de la DEPP, comme ayant importé leurs processus d'évaluation à travers PISA.

Non seulement la méthode américaine est jugée peu représentative et respectueuse des différents modèles éducatifs des pays participants, mais elle serait également réductrice dans la façon de présenter les résultats des évaluations, par des échelles aplatissant toute pluridimensionnalité analytique. L'auteur de la note s'inquiète de ces deux points :

« Les épreuves devant être adaptées à de nombreux pays sont, de ce fait, décontextualisées, c'est-à-dire indépendantes des programmes nationaux. Les performances des élèves aux exercices peuvent donc être très variables.

La présentation des résultats privilégie largement l'unidimensionalité (tous les pays placés sur une échelle unique) alors que l'univers des compétences serait plutôt multidimensionnel. »

On voit apparaître ici la différence de conception de la notion de performance telle qu'elle a été conceptualisée dans PISA, « ce que je sais faire avec ce que j'ai appris dans les situations scolaires » et celle appliquée dans les évaluations françaises, « ce que j'ai appris dans les situations scolaires »³⁵⁰. L'inquiétude exprimée est, d'une part, celle de voir la tradition évaluative française disparaître au profit de nouveaux outils d'évaluation jugés jusqu'alors inadaptés au système français, et d'autre part de ne pas détenir l'expertise adéquate pour utiliser ces nouveaux outils, moins bien maîtrisés par les évaluateurs nationaux. L'auteur de la note, en détaillant ses craintes, ne mentionne pas des aspects de méthode de PISA pourtant centraux dans le cadre de ses critiques, et en particulier de la technique de réponse à l'item employée dans PISA lorsqu'il précise que les performances des élèves peuvent être très variables selon les pays compte tenu des différences de programme. D'abord, PISA évalue la façon dont les élèves de 15 ans sont capables de se saisir des compétences, et non des connaissances, vues en classe pour répondre à des situations plus ou moins complexes issues de la vie quotidienne. Ensuite, les résultats ne sont pas construits à partir des réponses données par les élèves en tant que telles, puisque les niveaux des élèves par pays, et des établissements au sein de chaque pays, sont dérivés d'une série d'estimation de probabilités de réponse à des tests qui peuvent être différents. Les écarts de programme sont donc théoriquement gommés par la façon dont les items ont été d'abord construits, pour éviter tout biais notamment culturel, puis analysés pour être représentés sur une échelle de valeur. Enfin, la moitié des questions posées dans PISA mobilisent des connaissances vues au niveau primaire, et ce ne sont précisément pas ces connaissances qui sont l'objet de l'évaluation mais plutôt la façon dont elles sont mobilisées en autonomie par les élèves. L'un des agents de la DEPP, par ailleurs enseignant en mathématiques et ayant participé au cycle

³⁵⁰ Entretien avec Bruno Trosseille, Directeur du bureau des évaluations des élèves, DEPP, 14 avril 2016.

PISA 2012, exposait en entretien les implications de cette conception nouvelle en France dans le cas de la littératie mathématique sur le niveau recherché par les équipes de l'OCDE dans la création des items :

« Nous en France on était beaucoup plus sur un niveau de connaissances pures plus élevé, beaucoup plus exigeant. En termes de connaissances pures des connaissances d'algèbre, sur les fonctions, sur même en géométrie, mais avec des applications souvent directes de ces connaissances. Tandis qu'avec PISA on est sur des connaissances on va dire de base. On a calculé que la moitié des items PISA de 2012 sont basés sur des connaissances de primaire en France. En revanche, ce qui est plus souvent demandé aux élèves quand ils résolvent un exercice PISA, c'est de formaliser d'abord le problème mathématiquement à partir de ce qui peut être écrit, montré en image, dans un diagramme, dans un tableau, enfin bref encore une fois la prise d'information est souvent diverse, avec beaucoup d'écrit. Formuler mathématiquement et ensuite appliquer les connaissances qui sont d'un niveau assez faible. L'activité mathématique porte plutôt sur ce travail de formulation. On parle de modélisation, même si c'est discuté. Et ça c'est très peu enseigné en France.

Ce sont des choix conceptuels de l'OCDE qui sont construits, des choix qui viennent du concept de littératie mathématique, ou plutôt de littératie étendue aux mathématiques. On est dans une culture de l'écrit, et on va finalement évaluer la capacité des élèves à inférer sur ces écrits, et ça, ça va être la priorité. De manière annexe, et après avec des niveaux de difficulté de plus en plus grand évidemment, le fait que l'élève introduise ses connaissances et des connaissances de plus en plus abstraites ça va augmenter le niveau de difficulté de l'item et permettre de décrire ce que les élèves les plus performants sont capables de faire. Et ça c'est un choix conceptuel de l'OCDE. »³⁵¹

On voit avec cet exemple de littératie mathématique que le raisonnement mathématique surpasse le niveau d'enseignement des éléments de programme scolaire qu'il mobilise. La question du contenu différencié des programmes vu en fin de scolarité au collège entre les pays participants est donc peu pertinente dans le cas de PISA. L'éloignement de l'évaluation de l'acquisition d'un contenu de programme, impossible à réaliser dans des dizaines de pays aux programmes différents, a même été le point de départ de la construction de la méthode PISA

³⁵¹ Entretien avec Franck Salles, Chargé d'étude sur l'évaluation des élèves, National Project Manager, DEPP, 11 juillet 2017.

et la raison principale pour laquelle le choix a été porté sur l'évaluation d'élèves du même âge plutôt que d'un même niveau d'enseignement.

Les réserves exprimées par la DEPP au moment de la construction de la méthode de PISA portent ici sur des éléments classiques de la tradition évaluative française, qui a rejeté depuis près d'un siècle les principes de la psychométrie et cherche à asseoir la validité supérieure des méthodes qu'elle a conçues depuis plusieurs décennies. Elles se doublent d'une crainte de voir la formulation des questions, issue de traductions depuis l'anglais, accentuer les niveaux de difficulté pour les élèves français par rapport aux élèves d'autres pays, et notamment des pays anglophones.

d) Les méthodes de traduction des items : un désaccord de la DEPP classé sans suite

Alors que la DEPP possède en son sein des experts de la formulation des questions qui interviennent dans l'élaboration des évaluations maisons, la question de la traduction est en dehors de leur champ immédiat de compétence. S'ils proposent dans chaque cycle des items, et même si leur force de proposition a pu diminuer au fur et à mesure des années puisque la très grande majorité sont produits à partir des années 2010 directement par le consortium et seulement validés par les pays participants, ils ne sont pas chargés de leur traduction. Or, la traduction des items est centrale pour assurer la comparabilité des résultats obtenus dans une évaluation standardisée internationale, et pose des difficultés de prise en compte des spécificités linguistiques mais aussi culturelles de chaque pays participant³⁵². Les traducteurs doivent trouver des traductions strictement équivalentes d'une langue, et même d'un contexte linguistique à un autre (au sein des différents pays de la francophonie par exemple les mêmes formulations peuvent ne pas se référer aux mêmes situations quotidiennes), que cela soit en

³⁵² R.K. Hambleton et S Li, Translation and Adaptation Issues and Methods for Educational and Psychological Tests, dans C.L. Frisby et C.R. Reynolds (Eds.), *Comprehensive handbook of multicultural school psychology*, John Wiley & Sons, 2005.

termes de longueur de phrase, de fréquence d'usage, ou en degré d'abstraction nécessaire pour comprendre le sens du texte³⁵³. Alors que la standardisation tout au long du processus de passation des tests et de traitement des données implique, pour que les résultats soient robustes, d'avoir des méthodes similaires d'un pays à l'autre, la correspondance des méthodes doit dans le même temps tenir compte de la diversité des contextes. La langue étant porteuse d'un poids symbolique fort³⁵⁴, la maîtrise de la traduction devient dans le cadre de PISA un facteur important de la validité des résultats pour un contexte national donné. Or, le pouvoir de saisir en interne les inexactitudes de traduction et de faire évoluer les formulations des items n'appartient pas aux représentants nationaux – en l'occurrence la DEPP.

La note anonyme de la DEPP traite cette question de la traduction des items.

« Les méthodologies pour adapter (traduire) tests et questionnaires se sont largement développées depuis une vingtaine d'années. Néanmoins, ces méthodologies ne sont pas entièrement pertinentes dès lors que les épreuves portent essentiellement sur un matériel verbal complexe (cas de la littérature). On fait comme si la difficulté de compréhension d'un texte était la même dans toutes les langues.

Il faut également signaler que les méthodes d'identification des biais culturels ne permettent pas de les identifier tous lorsqu'on travaille sur des épreuves traduites car ces méthodes sont fiables pour travailler sur des preuves dans une même langue. »

L'assurance de l'équivalence linguistique porte en elle de nombreux défis auxquels la DEPP était jusque-là peu confrontée, et notamment celui d'intégrer non seulement une variété de situation d'apprentissage et d'enseignement, mais en plus de concevoir les évaluations en tenant compte d'une diversité de situations linguistiques et culturelles. Les problématiques de traduction dans des évaluations standardisées internationales avaient par ailleurs déjà surgi lors de

³⁵³ OCDE-PISA, *Consignes pour la traduction du matériel PISA*. Document provisoire préparé par Aletta Grisay, octobre 2998.

³⁵⁴ Jean-Marie Klinkenberg, *La langue et le citoyen. Pour une politique de la langue française*, Paris, Presses universitaires de France, 2001.

précédents travaux communs entre l'OCDE et la DEPP pour la France³⁵⁵, et en particulier IALS. Mais en tout état de cause, les mauvaises expériences d'un passé encore récent – IALS a été lancée en 1994 et les analyses françaises sur les causes de cet échec ont été rendues au début des années 2000³⁵⁶ – ont eu pour effet de laisser craindre des résultats moyens comparés aux autres pays participants, pouvant être expliqués notamment par des différences de niveau entre les items de différentes langues.

Les critiques portées à l'encontre de la traduction des items de PISA par, entre autres, la DEPP et les décideurs politiques de l'époque avaient été ainsi formulées de façon presque prémonitoire par T. Neville Postlethwaite, qui avait été consultant depuis le début des années 1990 pour la publication de l'OCDE *Regards sur l'éducation* au sein du réseau INES³⁵⁷:

« La façon traditionnelle d'entreprendre la traduction (des items) consistait à avoir deux traductions séparées et ensuite de les comparer. Là où il y avait un doute sur une phrase ou un item en particulier, les traducteurs (avec d'autres traducteurs si nécessaire), résolvaient le problème. Ensuite, une traduction dans le sens inverse était faite, et relue par ceux qui avaient originellement élaboré la version du test en anglais. Un grand soin était apporté à s'assurer que les mots employés étaient d'une difficulté similaire à ceux de la version originale. Là où on voyait des erreurs, elles étaient 'corrigées', et on anticipait que l'analyse des items après la phase de test montrerait d'autres problèmes dont on devrait s'occuper.

Le Programme PISA a demandé au groupe de Liège d'écrire des instructions détaillées (Document: TAG(98)_Translation Guide, et aussi la section dans le manuel destiné aux NPM). Les instructions données étaient exemplaires et étaient basées sur l'expérience de traduction des textes vers le français. Beaucoup de problèmes ont été identifiés dans ce premier processus de traduction. Ils ont été communiqués aux NPM. (Les instruments ont été

³⁵⁵ Ariane Baye, « La gestion des spécificités linguistiques et culturelles dans les évaluations internationales de la lecture », *Politiques d'éducation et de formation, analyses et comparaisons internationales*, 11/2004.

³⁵⁶ Voir notamment le travail de Alain Blum et France Guérin-Pace, *Des lettres et des chiffres : des tests d'intelligence à l'évaluation du « savoir lire », un siècle de polémiques*, Paris, Fayard, 2000, ou encore Marc Romainville, « Du bon usage de PISA », *La revue nouvelle*, vol. 115, issue 3-4, 2002, ou Gérard Bonnet et al., *The Use of National Reading Test for International Comparisons : Ways of Overcoming Cultural Bias*. Paris : Ministère de l'Éducation nationale, Direction de la Programmation et du Développement, 2001.

³⁵⁷ OCDE, DEELSA-PISA-BPC(2000) 3-Eng, p.33.

envoyés aussi bien en anglais qu'en français.) Les vérifications proposées paraissent tout à fait complètes, et je dis "chapeau".

Comme cela est déjà indiqué plus haut, le programme PISA porte en lui d'importants enjeux. Je peux imaginer qu'il y aura des critiques sur la comparabilité des difficultés de compréhension des mots utilisés. Les pays qui ne sont pas performants chercheront toujours des « excuses » pour expliquer pourquoi leurs élèves n'auront pas bien performés. Les NPM vont souhaiter se défendre et certains pourraient bien demander à PISA de réaliser la traduction pour eux pour qu'elle soit "vue comme juste". Ce sera très difficile pour PISA de s'organiser. La traduction des tests et des items des questionnaires n'est pas un travail "ordinaire" de traduction comme il peut être fait par des agences de traduction professionnelles. Il me paraît essentiel que quelqu'un du Centre National ou du comité d'expert soit impliqué dans les processus à la fois de traduction et de réconciliation. »³⁵⁸

T. Neville Postlethwaite s'exprime ainsi dans un rapport réalisé avec Jeanne Griffith et Hans Günter Rolff pour le GPP en 2000, visant à examiner en quoi PISA pourrait permettre de fournir des données utiles et valides pour être utilisées à des fins d'orientation de politique éducative, en quoi le processus développé par l'OCDE et le GPP serait efficace en temps et en coût, en quoi il permettrait des

³⁵⁸ "The traditional way of undertaking the translation was to have two separate independent translations and then compare them. Where there was doubt about a particular phrase or item, the translators (with others if necessary) would resolve the problem. Then a separate back translation would be made and checked by those responsible for the original English version of the test. Care had to be taken to use words of similar difficulty levels to the original. Where there were glaring errors these were then 'put right', and it was expected that the item analysis after the trial testing would disclose further problems that would then be 'mended'.

The PISA programme asked the Liège group to write detailed instructions (Document: TAG(98)_Translation Guide, and also the section in the NPM manual). The instructions were exemplary and were based on the experience of already having translated the materials into French. Many problems were identified in this initial translation process. These were communicated to the NPMs. (The instruments were sent out in both English and French.) The proposed checking seems to be very thorough. 'Chapeau!' is my reaction.

As already indicated above, the PISA programme is a high-stakes enterprise. I can imagine that there may well be criticisms of the comparability of the difficulty level of words used. Countries that do not perform well will always look for 'excuses' as to why their pupils did not perform so well. The NPMs will wish to defend themselves and some may well ask PISA to undertake the translation for them so that the translation is 'see to be fair'. This will be very difficult for PISA to organise. The translation of test and questionnaire items is not an 'ordinary' translation job as usually conducted by professional translation agencies. It would appear to me to be essential that someone from the National Centre or from the expert committee be involved in the translation and reconciliation process." Traduit par nous.

résultats équilibrés entre ce qui relève de la mesure statistique et ce qui relève de l'analyse de politique publique et en quoi enfin l'évaluation permettrait aussi bien de tirer des conclusions solides au niveau international et national.

Les alertes qu'il émet et les critiques qu'il prévoit sont partagées par un large pan de la recherche universitaire sur l'effet de la traduction sur la performance des élèves à une évaluation³⁵⁹. Les solutions dont il salue la mise en œuvre par l'OCDE, et notamment la sollicitation d'un groupe de travail d'experts renommés, habitués par ailleurs des traductions vers le français, aboutiront à la publication précise de la méthode de double traduction employée dans le cadre de PISA, pour en assurer l'accessibilité et la compréhension par tous. Le document de « *Consignes pour la traduction du matériel PISA* ³⁶⁰», publié par l'OCDE en 1998, détaille un dispositif unique par son ampleur, avec une multiplication forte des contrôles effectués sur la qualité des traductions, et tout cela avant la phase de pré-test qui permet elle aussi d'éliminer encore les erreurs non détectées auparavant. Deux traducteurs professionnels travaillent d'abord indépendamment l'un de l'autre avant que leurs traductions ne soient réunies par un troisième traducteur, souvent non professionnel mais au profil de chercheur et fin connaisseur des disciplines évaluées par les items traduits. Une équipe constituée de traducteurs de plusieurs pays vérifie ensuite les traductions, avant qu'elles ne soient testées. Le matériau de base est par ailleurs fourni aux pays en anglais et en français pour que les traducteurs puissent s'appuyer sur deux langues pour effectuer leurs traductions. La limite reste que la version française « de base » est en fait le plus souvent traduite de l'anglais, même si le processus reste le même que celui utilisé pour toutes les autres traductions de PISA.

L'attention demandée par la DEPP pour effectuer les traductions après le raté de IALS semble avoir été entendue par l'organisation internationale pour PISA. L'une de nos enquêtées, analyste à l'OCDE pour PISA depuis les débuts de l'évaluation, indiquait ainsi en entretien que ces recommandations de la

³⁵⁹ Ariane Baye, *art. cit.* p.2

³⁶⁰ OCDE-PISA, *Consignes pour la traduction du matériel PISA*, 1998, document préparé par Aletta Grisay.

France avaient été suivies des faits, et notamment en termes d'investissement de la part de l'OCDE, et donc des contributions demandées aux pays participants.

« Des budgets énormes sont déployés pour la traduction des questionnaires afin d'éviter les biais culturels. C'est vraiment central dans le travail qu'on fait, vu qu'on parle de littératie et donc de compréhension à la lecture. »³⁶¹

Alors que la DEPP semble avoir eu gain de légitimité auprès de l'OCDE sur le sujet des traductions d'items, les inquiétudes au sujet de l'équivalence entre les langues et les contextes culturels sont tout de même formulées par les agents au moment des premiers résultats de PISA. Ces critiques évoluent toutefois assez rapidement pour laisser place à d'autres réactions. En effet, progressivement, et dans le même mouvement d'acceptation de la méthode psychométrique peu usuelle pour les évaluateurs français, les critiques portant sur les risques liés à une traduction des items de mauvaise qualité laissent place à des réactions de tri de ce qui est exploitable pour orienter la politique éducative française.

« Le plus intéressant, c'est quand même l'attitude franco française, que j'ai moins vécue, avec les premiers PISA qui ont été considérés en France comme un machin construit par les anglo-saxons, donc ça ne peut pas s'appliquer à nous. Les résultats sont faux parce qu'on est différents. PISA c'est un truc très sérieux, il ne faut pas lui faire dire ce qu'il ne dit pas mais il n'y a rien à dire sur la méthode. »³⁶²

« Je crois qu'on est toujours sur une part de 70% d'items d'origine anglo-saxonne. Ce qui est drôle c'est qu'on avait fait ces analyses là avec Thierry en 2000, et on se rendait compte que les élèves français sont meilleurs sur les supports français. C'est quand même marrant, donc oui, forcément ça joue. Donc forcément on peut supposer que lorsque les items sont majoritairement en provenance d'autres pays on les réussit moins bien. Et so what ? Enfin je veux dire on ne va pas aller bien loin avec ça. Mais c'est quand même pour ça je crois qu'il faut vraiment regarder ce que ça nous apprend sur nous. Après, si on accorde trop d'importance à cette idée de provenance d'items et qu'on réussit moins bien, qu'il y a des biais culturels, de traduction, tout ça, on ne fait plus rien ou on n'avance pas. Il

³⁶¹ Entretien avec Sophie Vayssettes, responsable de projet PISA, OCDE, 5 novembre 2015.

³⁶² Entretien avec Catherine Moisan, Directrice de la DEPP de 2013 à 2015, 10 juin 2020.

faut le savoir mais en même temps il ne faut pas en faire un argument, ça me semble assez stérile. »³⁶³

L'influence de la DEPP n'a pas été suffisante pour aboutir à une plus grande sollicitation des NPM dans la construction des items, et le constat reste celui de problèmes d'équivalence des niveaux de difficulté des items d'une langue à l'autre combinés à une prédominance des items jugés « anglo-saxons », soit parce qu'ils sont fournis par les pays anglo-saxons de PISA, soit parce qu'ils sont créés par des groupes de concepteurs du consortium à dominante anglo-saxonne.

Néanmoins, le choix est fait par les évaluateurs nationaux de changer de prisme d'analyse des résultats PISA pour la France en se détachant d'une critique méthodologique sur laquelle ils n'ont pas pu avoir de prise satisfaisante, jugée en interne comme « *un mal nécessaire* ». Certes les modèles d'évaluation sont différents entre la tradition française dont la DEPP est l'héritière et celle de la psychométrie considérée comme anglo-saxonne, mais il ne s'agit plus d'ériger ces différences en oppositions puisque la méthode est jugée rapidement solide par la plupart des évaluateurs des pays participants, et que rien ne la fera désormais évoluer vers davantage de prise en compte de l'approche française. Il faut plutôt, après quelques années, imaginer un usage valorisant le modèle d'évaluation français.

2) Faire de PISA un faire-valoir des évaluations nationales de la DEPP

Les critiques méthodologiques fortes formulées par les évaluateurs nationaux après la sortie des premiers résultats PISA et celles émises en interne résumées dans la note anonyme dont nous avons retranscrit ici la plus grande part, n'ont pas donné lieu cette fois, comme pour IALS, à une proposition de retrait de l'évaluation. PISA a été plutôt considéré par les évaluateurs comme inévitable :

³⁶³ Entretien avec Isabelle Robin, Directrice du département de la recherche et du développement, de l'innovation et de l'expérimentation, Dgesco, 7 avril 2017.

il semblait à leurs yeux impensable de se retirer de PISA. Si la perception de la contrainte semble claire, elle s'accompagne (et est en partie renforcée) par la volonté de s'approprier le dispositif, à la fois pour en maîtriser les dangers et pour le mettre au service de l'expertise nationale.

Le choix de la DEPP a dès lors consisté à assurer une présence très active dans les réunions techniques, à l'inverse d'une politique de la chaise vide. Elle a en parallèle cherché à tirer avantage de la médiatisation croissante de PISA pour créer de nouveaux outils d'évaluation susceptibles d'intéresser davantage les décideurs publics, au premier rang desquels l'évaluation CEDRE, et pour valoriser ses propres activités d'évaluation auprès des acteurs de la politique éducative française. D'une certaine manière il s'agissait à la fois de s'accommoder de PISA et de l'utiliser pour promouvoir d'autres instruments d'évaluation.

Pour cela, le débat sur l'utilité de PISA est progressivement repositionné par la DEPP sur le niveau national, pour faire de l'évaluation un révélateur des problématiques relatives au système éducatif français comparé à lui-même dans le temps, en dehors de toute logique de classement international. Ce resserrement du cadre spatial de PISA a été rendu possible une fois plusieurs cycles passés, ce qui a permis de comparer les résultats obtenus par les élèves français sur une même discipline majeure à quelques années d'intervalle.

a) La comparabilité dans le temps : une évaluation internationale nationale

Au sein de la DEPP, l'argument de comparabilité internationale permise par l'évaluation telle qu'elle est menée par l'OCDE est rapidement ramené à un argument politique impliquant l'emploi d'une bonne partie du budget de la direction. Le vrai usage des résultats résiderait davantage dans la comparabilité des données relatives aux élèves français dans le temps. L'un de nos interlocuteurs, ancien NPM, le formulait ainsi, sans détour :

« PISA est fait de façon sérieuse. Ça permet de faire un suivi dans le temps de nos élèves, et ça c'est intéressant. On pourrait le faire pour 10 fois moins cher en interne mais ce sont des choix politiques. »³⁶⁴

L'intérêt de PISA, une fois actée sa dimension politique plus que témoignant de manques dans les évaluations nationales, est à trouver dans ce qu'il permet de comprendre de soi et est donc à chercher dans des comparaisons temporelles de ses propres résultats. Cela est rendu possible au bout de neuf ans et le quatrième cycle PISA qui reprend la discipline majeure du cycle initial de 2000. De la même façon que pour CEDRE les évaluateurs de la DEPP jugent que les premiers résultats vraiment intéressants sont ceux des évaluations postérieures à 2009, pour avoir une comparaison possible avec celles menées six ans plus tôt et donc le recul d'une génération d'élèves ayant vécu les réformes mises en place avant leur entrée au collège, les résultats de PISA se voient conférer une valeur similaire à une évaluation nationale par échantillon après le cycle 2006.

Le mot d'ordre des évaluateurs de la DEPP devient celui de l'analyse de la « trajectoire » des résultats des élèves français et pas d'un « palmarès » des pays. Les entretiens que nous avons réalisés avec des évaluateurs de la DEPP en poste à différents moments depuis le lancement de PISA témoignent bien de ce changement de posture.

« L'évaluation PISA est légitime, sauf que ce n'est pas une affaire de palmarès mais plutôt de trajectoire. C'est plus ce que ça nous apprend sur nous-même. »³⁶⁵

« La partie analyse c'est le plus intéressant, mais on a très peu de temps. On a des groupes d'analyse qui sont coordonnés par des gens en interne à la DEPP. On fait des réunions régulières pour analyser dans le temps les

³⁶⁴ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 21 mars 2019.

³⁶⁵ Entretien avec Ronan Vourc'h, Adjoint au chef de bureau de l'évaluation des élèves, DEPP, 22 juin 2020.

résultats de nos élèves, voir s'il y a une optique particulière. On ne regarde pas les résultats des autres pays. »³⁶⁶

« Alors c'est quand même une évaluation qui est parfaitement légitime. Je ne mets pas en doute la légitimité de PISA. Sauf qu'un petit peu comme d'habitude, c'est évidemment pas une affaire de palmarès. Je pense qu'il n'est pas très intéressant de savoir derrière qui ou devant qui on est. Enfin bon, ce n'est pas inintéressant non plus parce qu'on a des trajectoires. Mais le plus intéressant dans PISA c'est quand même ce que ça nous apprend sur nous-même. C'est ça, ce n'est pas la comparaison. »³⁶⁷

L'évaluation PISA devient cette fois légitimée par les agents de la DEPP, en tant qu'évaluation ayant un sens au niveau national, bien qu'ils n'aient pas complètement le contrôle de sa méthode. L'idée n'est pas de voir les résultats des élèves français à l'aune de l'ensemble des résultats des autres participants. Elle n'est pas non plus de voir en quoi l'évolution des résultats français serait proche ou différente de l'évolution des résultats des autres pays dans le temps. Il s'agit désormais, dès 2009, de regarder ce que PISA dit du système français par rapport à lui seul et à chaque fois que des disciplines comparables sont évaluées sur une population aux caractéristiques similaires et à des moments différents.

Le repositionnement sur l'évolution des résultats nationaux présente plusieurs intérêts dans la construction de l'identité d'expertise de la DEPP dans le champ de l'évaluation de la politique éducative. D'abord, elle permet aux agents de la DEPP de produire une analyse des résultats qui nourrit des réflexions de politique éducative, et donne, par les thèmes qu'ils privilégient, des propositions d'orientation pour l'action publique. Le temps dont ils disposent pour analyser les résultats PISA dans leur ensemble est particulièrement réduit compte tenu des rythmes de publication de l'OCDE et de la petitesse des équipes qui y travaillent à la DEPP. Si pendant la décennie 2000 les productions de la DEPP sur PISA se sont principalement concentrées dans les notes d'information synthétiques donnant à voir de façon descriptive les résultats jugés les plus

³⁶⁶ Entretien avec Anne-Laure Monnier, Chargée de mission au Conseil supérieur des programmes, 12 septembre 2016.

³⁶⁷ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 18 mars 2019.

saillants pour les élèves français, les années 2010 voient se multiplier les articles et les communications dans des colloques et congrès autour de problématiques nourries par l'évolution des résultats de la France à PISA. La comparaison des résultats français à PISA de cycles à d'autres est ainsi utilisée, par exemple, pour mettre en avant les conclusions de la DEPP sur l'effet contre-productif du redoublement³⁶⁸, sujet qu'elle tentait de mettre à l'ordre du jour des réflexions des décideurs politiques depuis des années sans succès. Elle a également permis à la DEPP de fournir des données propres et des analyses sur les facteurs aggravant les inégalités scolaires, comme la mixité dans les établissements ou encore l'effet de l'organisation scolaire sur la réduction des inégalités sociales.

Second intérêt pour le positionnement de la DEPP dans le champ de l'expertise en politique éducative, le changement d'envergure de l'analyse des résultats de PISA permet de renouer des contacts plus forts avec les acteurs centraux de l'Education nationale, les enseignants, qui sont peu investis dans PISA et n'en reçoivent que peu d'informations venant d'une communication spécifique de leur ministère de tutelle. Catherine Moisan, directrice de la DEPP de janvier 2013 à novembre 2015, illustre en entretien cette évolution aussi bien du discours que des usages des résultats français à PISA. Elle estimait ainsi que PISA était une vraie étude nationale, plus qu'un outil de comparaison internationale. Elle avait initié en conséquence un travail de mise en valeur du travail de la DEPP auprès des enseignants par l'intermédiaire des écoles de formation, en combinant les résultats des élèves français à PISA et aux évaluations nationales sortis la même année.

« Je préside le conseil de l'ESPE de Créteil. J'ai fait l'an dernier une conférence pour les formateurs de l'ESPE, je leur ai raconté plein de trucs intéressants sur PISA et les panels, pour la première on a eu des résultats en 6e et en 3e, les mêmes élèves, Thierry (Rocher) avait réussi à faire une échelle pour comparer, et c'était extrêmement percutant sur quatre ou cinq compétences. Il donnait des exemples, et les formateurs découvraient le sujet. C'est un sujet qui m'importe beaucoup, et c'est difficile, c'est que ça touche les profs. »³⁶⁹

³⁶⁸ Thierry Rocher, Comparaisons internationales, *Education & Formation* n°78 numéro spécial, novembre 2008.

³⁶⁹ Entretien avec Catherine Moisan, Directrice de la DEPP de 2013 à 2015, 10 juin 2020.

Afin de partager au plus grand nombre d'enseignants les résultats les plus forts, qualifiés même de « percutants » sur certaines compétences, la porte d'entrée choisie avait été celle de renforcer la communication sur les résultats de PISA avec leurs formateurs, ce qui se faisait peu auparavant. L'absence de valorisation volontaire des classements de PISA et de la position de la France dans ceux-ci, et le lien fait avec les travaux de la direction au niveau national a permis de faciliter l'écoute de formateurs dont le rôle est d'accompagner les enseignants actuels et à venir dans l'évolution de leur pratique au quotidien. Mieux à mêmes de réutiliser dans leurs formations les résultats présentés comme nationaux plutôt qu'internationaux, et les cas types présentés par la DEPP lors des conférences dédiées aux évaluations des résultats des élèves, les formateurs ont pu faciliter la transmission d'information de la DEPP jusqu'aux enseignants.

En faisant en sorte de faire circuler les résultats de PISA jusqu'aux enseignants, la DEPP sous la direction de Catherine Moisan renoue sous une certaine forme avec le souhait de Claude Thélot d'un « effet miroir » des évaluations dans les établissements scolaires pour accompagner l'évolution des pratiques d'enseignement. Cet « effet miroir », qui pour Thélot consistait à donner à voir aux enseignants les façons d'apprendre et d'enseigner dans d'autres contextes d'apprentissage³⁷⁰, ne se fonde pas ici comme il l'imaginait sur des contextes différents par le pays dans lequel ils s'inscrivent mais dans des temporalités différentes.

On le voit apparaître en filigrane dans les exemples cités, le changement de discours de la DEPP, et en particulier l'étiollement d'une stratégie initiale de critique ouverte de la méthode, reflète à la fois une contrainte perçue (ils pensent ne pas pouvoir échapper à PISA) mais aussi la volonté de se réappropriier le dispositif PISA pour repositionner au premier plan leur expertise en matière d'accompagnement de l'action publique en éducation. Cette stratégie est rendue

³⁷⁰ Xavier Pons, *Evaluer l'action éducative. Des professionnels en concurrence*, Paris, PUF, 2010.

possible grâce au caractère cyclique de PISA et aux comparaisons nationales d'une décennie à l'autre qu'il a permis après quelques années. Trois dimensions de cette stratégie ont été particulièrement mises en avant lors de nos entretiens et lectures et renseignent sur l'évolution de la place des évaluateurs nationaux de la DEPP au sein des acteurs de l'évaluation en politique éducative : la création d'un nouvel outil, la constitution d'un discours de complémentarité des données, et le rapprochement des travaux de la DEPP du monde de la recherche.

D'abord, la DEPP s'est dotée d'un outil supplémentaire, l'évaluation CEDRE, très rapidement après le lancement de PISA, convaincue de la puissance de la conception nouvelle des résultats de l'éducation adoptée par l'OCDE pour élaborer son évaluation. Par puissance, la DEPP entend d'une part l'intérêt prospectif du choix de se centrer sur la littératie pour construire les items d'évaluation, qu'il s'agisse de celle de la maîtrise de la langue, une application déjà familière pour les évaluateurs nationaux même si peu utilisée dans les évaluations qu'ils menaient jusque-là, ou de celle des compétences dans les disciplines scientifiques. Elle attend d'autre part une écoute attentive de la part des décideurs politiques nationaux.

L'évaluation CEDRE est la première pierre posée d'un dispositif visant la valorisation de la complémentarité des données et de l'apport de PISA pour confirmer les résultats des évaluations nationales menées toutes par la DEPP.

Ce discours de complémentarité des données, qui implique la maîtrise technique de plusieurs traditions évaluatives par les agents de la DEPP qui le tiennent, sera tenu non seulement auprès des décideurs politiques, premier interlocuteur de la DEPP dans ses missions de service statistique ministériel, mais aussi auprès du monde de la recherche dont elle se sent proche, au moyen d'une revue à comité de lecture qu'elle édite, *Education & formations*.

b) PISA, un brouillon pour CEDRE : la reprise en main de l'expertise méthodologique par les évaluateurs nationaux

Très rapidement après le lancement et la passation du cycle test du premier PISA en 2000-2001, la DEPP commence à travailler sur une évaluation « maison » qui en reprend certains principes clés. En 2003 est lancé ce nouveau produit de la DEPP nommé le « cycle des évaluations disciplinaires réalisées sur échantillon », plus communément connu sous son acronyme CEDRE. L'évaluation s'adresse aux élèves en fin d'école primaire âgés donc de 9-10 ans, et en fin de collège, âgés de 14-15 ans. Elle est fondée sur une approche d'évaluation des compétences en conservant la focale disciplinaire à laquelle la DEPP était habituée par ses autres évaluations, et l'objectif d'un apprentissage de programmes d'enseignement par discipline³⁷¹. En plus de ce premier sujet d'évaluation, d'autres approches se combinent pour donner des éléments d'évaluation sur des aspects non disciplinaire et transdisciplinaires acquis en classe. CEDRE analyse, ainsi par exemple, la créativité des élèves, la façon dont ils collaborent entre eux, leurs façons de communiquer. Le dispositif interroge également le niveau de maîtrise de la langue à travers des questions recoupant plusieurs disciplines. On retrouve dans ces deux dernières approches des objets d'évaluation popularisés par PISA, qu'il s'agisse de la littératie des élèves, proche de la maîtrise transdisciplinaire de la langue, ou de l'évaluation de paramètres individuels n'ayant pas trait à du contenu scolaire mais à l'attitude des élèves dans une situation de vie en société.

Le processus de construction des tests et des items fait intervenir bien sûr les agents de la DEPP, et notamment ceux qui interviennent aussi sur PISA, accompagnés d'enseignants recrutés pour l'occasion, d'inspecteurs de l'éducation nationale et de formateurs³⁷². Cette équipe composée de plusieurs profils est

³⁷¹ Voir par exemple la synthèse réalisée par l'Inspection générale de l'éducation nationale, *Les évaluations nationales et internationales sur les acquis des élèves et sur d'autres dimensions des systèmes éducatifs*, rapport n°2012-056, mai 2012.

³⁷² Brigitte Grugeon-Allys, Nadine Grapin, « Validité d'une évaluation externe. Complémentarité des approches didactique et psychométrique ». *Actes du séminaire national de l'Association pour la Recherche en Didactique des Mathématiques*, mars 2015.

importante pour l'objectif qu'assigne la DEPP à cette évaluation : observer les performances des élèves en tenant compte à la fois des programmes supposés vus par les élèves qui passent le test, les attentes officielles de l'éducation nationale et les instructions fournies aux équipes enseignantes et les pratiques d'enseignement en classe³⁷³.

Par cette création interne d'une évaluation nationale standardisée, la direction se dote d'un outil qui lui permet de reprendre en main la maîtrise de l'expertise en matière d'évaluation du système éducatif français. L'évaluation ne provient pas d'une commande ministérielle au sens strict, comme ont pu l'être à différentes périodes les évaluations-diagnostic exhaustives menées auprès de l'ensemble des élèves de plusieurs niveaux scolaires. Elle est construite à l'initiative des agents de la DEPP, même si elle s'inscrit dans un contexte plus général où différentes instances évaluatrices de l'Education Nationale, l'Inspection générale et le Haut Conseil de l'Evaluation de l'Ecole (HCée) soutiennent l'idée d'un tel projet.

Plusieurs membres de l'Inspection Générale de l'éducation nationale attribuent la naissance de CEDRE aux recommandations formulées par le Haut Conseil de l'Evaluation de l'Ecole en 2001³⁷⁴ dans le rapport sur *Les forces et les faiblesses de l'évaluation du système éducatif français*³⁷⁵, notamment pour compléter l'arsenal d'évaluations déjà menées par la DEPP avec des évaluations permettant des comparaisons dans le temps. Ils en rendent compte notamment dans un rapport exhaustif sur les évaluations nationales et internationales à la fois des résultats des élèves et des performances du système éducatif, qu'ils remettent au ministre de l'Education nationale en 2012. Les évaluations standardisées sont bien évoquées largement dans le rapport du HCée pour inciter à s'intéresser aux évaluations des élèves sur échantillon comme outil d'évaluation du système

³⁷³ Thierry Rocher, « Mesure des compétences. Méthodes psychométriques utilisées dans le cadre des évaluations des élèves », *Education & Formations*, n°86-87, mai 2015.

³⁷⁴ Jean-Paul Delahaye, Marie-Blanche Mauhourat, François Monnantueil, Yves Poncelet, Anne Vibert, Johan Yebbou, Inspection générale de l'éducation nationale, *Les évaluations nationales et internationales sur les acquis des élèves et sur d'autres dimensions des systèmes éducatifs*, rapport à Monsieur le ministre de l'éducation nationale, mai 2012, n°2012-056

³⁷⁵ Claude Pair, *Les forces et les faiblesses de l'évaluation du système éducatif français*, HCée, 2001.

éducatif dans son ensemble³⁷⁶. Si elles ne sont pas rattachées explicitement à PISA et si l'OCDE n'est pas nommément citée, tous les principes fondateurs et les débats ayant précédés le lancement de cette évaluation internationale spécifique sont bien présents dans la description des besoins faite par le HCée.

« L'évaluation peut avoir un autre but, non plus individuel mais global : rendre compte des résultats du système éducatif par un bilan des connaissances et compétences acquises par l'ensemble d'une classe d'âge, ou par l'ensemble des élèves à certains niveaux ; et si possible envisager son évolution dans le temps. Les examens ne se prêtent pas à cet usage. Il faut plutôt utiliser des épreuves standardisées appliquées à un échantillon représentatif de la population considérée. (...) La validité de ces « évaluations-bilans » pose des questions d'ordre psychométriques (Salines et Vrignaud, 2001) ; en particulier la question du choix du référent est délicate. Faut-il se rapporter aux programmes scolaires, alors qu'il n'est pas certain qu'ils reflètent bien les besoins sociaux et alors qu'ils sont fréquemment modifiés ? Ou, comme pour certains tests internationaux, tenter de définir ces besoins, mais se mettre d'accord sur eux est bien difficile et eux-mêmes varient dans le temps ? »³⁷⁷

On voit là que le principe qui fait toute l'originalité de PISA, l'évaluation selon l'utilisation des compétences acquises et plus selon les programmes censés être vus par les élèves, est présenté comme l'une des démarches d'évaluation possibles, sans pour autant encore une fois que l'étude ne soit nommée par l'auteur du rapport. Est-ce lié à une volonté de ne pas articuler les recommandations du HCée aux activités d'une organisation internationale aux préceptes libéraux, et récemment entrée dans le champ des évaluations des politiques éducatives ? Ou tout simplement au fait qu'en 2001, à l'époque où le rapport du HCée est travaillé et publié, le premier cycle de PISA n'est pas encore terminé et que son acronyme n'aurait que peu de sens aux yeux des lecteurs ? S'il est difficile de répondre à cette question avec les éléments dont nous disposons, il est en revanche plus que probable que l'évaluation évoquée par Claude Pair soit bien celle de l'OCDE. Outre la mention des principes psychométriques, de

³⁷⁶ Claude Pair, *op. cit.* p.17

³⁷⁷ Claude Pair, *op. cit.* p.17.

l'évaluation prospective des besoins sociaux comme moyen de construire les résultats attendus de l'éducation, et bien sûr celle de « tests internationaux », le HCcéé donne également de sérieux indices sur l'évaluation internationale dont il est question. En effet, si Claude Pair, ancien recteur d'académie, est l'auteur du rapport, Claude Thélot en est alors le directeur de publication au titre de la présidence du HCcéé qu'il occupe. Tout juste passé de la direction de la DEPP, qu'il a occupée pendant sept ans de 1990 à 1997, à celle du HCcéé, il n'a jamais caché son intérêt pour PISA et pour les évaluations standardisées, et a été à l'initiative de l'implication de la DEPP dès le début de la construction de PISA par l'OCDE et les pays membres.

Au sein de la DEPP, l'inspiration de PISA pour le lancement de CEDRE a été affirmée par plusieurs de nos interlocuteurs en entretien et même revendiquée comme ayant permis de renforcer la validité scientifique de la méthode employée et, par dérivation, des résultats obtenus. PISA aurait été en quelque sorte un « brouillon », un galop d'essai méthodologique pour repérer les points de méthodes innovants, légitimes aux yeux des décideurs politiques, et apportant quelque chose à l'arsenal déjà existant d'évaluations disponibles en interne.

Le DEPP reprend donc ce qui lui paraissait utile de PISA pour affirmer sa capacité à innover dans ses pratiques d'évaluations et à retrouver une identité d'entrepreneur de problème public auprès de la communauté éducative et de la classe politique. Les principes de la psychométrie, pourtant très décriés dans le champ des évaluations des résultats de l'éducation en France depuis des décennies, font avec CEDRE leur entrée au sein du panel d'évaluation de la DEPP. Comme l'indique en entretien Bruno Trosseille, la « philosophie » derrière CEDRE est la même que celle de PISA.

« Les procédures étaient issues de PISA, avec des cahiers tournants, des modèles de réponse à items, des modèles de Rasch à un paramètre, ou deux paramètres. La philosophie est la même. »³⁷⁸

³⁷⁸ Entretien avec Bruno Trosseille, Directeur du bureau des évaluations des élèves, DEPP, 14 avril 2016.

Une membre de l'équipe de statisticiens en charge de PISA pour la DEPP nous dit encore à propos de la méthode employée dans PISA :

« La méthode est intéressante, on s'inspirait beaucoup de PISA pour les autres évaluations ; et là on a le modèle d'évaluation par item qu'on utilise aussi. Nous on a une évaluation qui s'appelle CEDRE et qui a beaucoup de similitudes. »³⁷⁹

Le modèle de réponse à l'item (MRI) est réellement une nouveauté pour la DEPP directement issue de PISA³⁸⁰. Il consiste à sortir d'un modèle de test classique pour permettre d'établir des scores individuels sur une même échelle de compétences pour des élèves différents qui ont reçus des items différents et ont donc *a fortiori* répondu à des items différents³⁸¹. On part du principe que la réponse à un item est liée à la fois à des caractéristiques de l'item en lui-même et aux caractéristiques de l'individu qui y répond. Cela permet schématiquement de pouvoir estimer la difficulté d'un item, des probabilités de réponse correcte par hasard et de donner, indépendamment de la constitution de l'échantillon de répondants, un niveau de compétence pour chaque répondant dit « latent », c'est-à-dire non mesurable immédiatement. En d'autres termes, ce n'est pas parce qu'on ne donne pas une réponse correcte à un instant t qu'on ne dispose pas des compétences permettant d'y répondre et ce n'est pas parce qu'on donne une réponse correcte à un instant t qu'on en dispose, à l'échelle d'un échantillon de répondants dans un contexte donné. On sort, avec cette méthode, des pratiques habituelles de la direction qui, dans la lignée de la tradition évaluative française

³⁷⁹ Entretien avec une statisticienne, en charge de PISA, DEPP, 1er juillet 2020.

³⁸⁰ Sébastien Béland, David Magis et Gilles Raïche, Estimation des paramètres d'item et de sujet à partir du modèle de Rasch : une étude comparative des logiciels BILOG-MG, ICL et R. *Mesure et évaluation en éducation*, 36(1), 2013.

³⁸¹ Voir par exemple Dany Laveault et Jacques Grégoire, *Introduction aux théories des tests en psychologie et en sciences de l'éducation*, Bruxelles, De Boeck, 2002 ; Frank B. Baker et Seock-Ho Kim, *Item Response Theory – Parameter Estimation Techniques*, 2^e ed. New York, Marcel Dekker, 2004 ; Richard Bertrand, Jean-Guy Blais, *Modèles de mesure. L'apport de la théorie des réponses aux items*, Sainte-Foy, Presses de l'Université du Québec, 2004.

construite au long du XXe siècle, s'était écartée des méthodes d'évaluation issues de la psychométrie, courant pourtant initié par le Français Alfred Binet³⁸².

Autre aspect de cette « philosophie » commune entre CEDRE et PISA, le principe de temporalité de l'évaluation est repris, ainsi que les différentes étapes de validation des cycles à venir. La DEPP a en effet choisi de reprendre le principe de renouvellement des évaluations pour une même discipline à intervalles réguliers pour pouvoir comparer l'évolution dans le temps. La première comparaison possible a eu lieu en 2009 sur la même discipline que la première évaluation menée en 2003, sur la maîtrise de la langue pour ce qui est des élèves de CM2 et les compétences générales et langagières pour les élèves de 3^e. L'évaluation comporte, comme PISA, une phase de test organisée avant la passation du test officiel, pour permettre une plus grande validité de l'analyse des résultats en écartant après le test les items ne correspondant pas aux exigences statistiques du modèle de réponse à l'item. Une statisticienne en charge de PISA au sein de la DEPP nous précisait en entretien :

« On fait une phase de préparation aussi, on fait un test avant l'évaluation. C'est important et on est tous les ans en train de préparer l'évaluation et de passer les tests. CEDRE, on est sur toutes les matières. On a histoire-géo, sciences, des choses qui ne sont pas forcément sur PISA. Ce qui est bien et qui se passe très bien c'est qu'on travaille tous ensemble, on fait des réunions, on s'échange des informations. »³⁸³

La multiplicité des matières évaluées a entraîné une hausse de la masse de travail pour les équipes de la DEPP. La proximité méthodologique avec PISA a permis, même si les méthodes psychométriques étaient nouvelles pour l'ensemble de l'équipe, de gagner en maîtrise des processus de passation et de construction des tests. Il est intéressant de noter que dès 2002 l'ensemble de l'équipe en charge de CEDRE a travaillé à la construction des cycles en lien direct avec l'équipe en charge de PISA pour la DEPP, sans que cela ne soit source de critiques méthodologiques fortes, ou en tout cas pas suffisamment fortes pour remettre en

³⁸² Voir chapitre 1 pour plus de détails.

³⁸³ Entretien avec une statisticienne, en charge de PISA, DEPP, 1er juillet 2020.

cause l'orientation choisie par le bureau de l'évaluation des élèves dans sa nouvelle évaluation « maison ». Si l'analyse critique de la méthodologie de PISA est forte au moment où sort le premier cycle PISA, et notamment venant des évaluateurs de la DEPP, la reprise en interne de ces mêmes principes méthodologiques paradigmatiques du QCM et du modèle de réponse à l'item, pose peu, voire pas, question. Il est même pris en modèle par les équipes de concepteurs de CEDRE et par les analystes de la DEPP qui tireront des conclusions comparées entre les résultats d'items proches à PISA et à CEDRE. Certains items de PISA sont ainsi utilisés et revendiqués comme étant une vraie inspiration pour les items de CEDRE, comme cela a pu nous être indiqué en entretien.

« On peut s'inspirer pour CEDRE, on n'a pas le droit de reprendre des items, mais bon on s'inspire, on fait des liens. CEDRE c'est sur la classe de 3^e, les résultats on les met en lien. C'est pas la même chose, pas le même construit, PISA c'est des questions assez générales qui mettent à profit des compétences dans un contexte de vie courante. Nous on s'appuie sur des questions scolaires. »³⁸⁴

L'inspiration de PISA n'est pas seulement une façon de gagner du temps, mais bien de « faire des liens » entre les évaluations portant sur les élèves de 15 ans pour PISA et de 3^e pour CEDRE. L'usage est ouvertement analytique, visant à outiller l'action publique, et il ne s'agit pas de vérifier d'éventuels écarts de résultats explicables par des problèmes de méthode de passation ou de construction des échelles de résultat. La comparaison stricte n'est évidemment pas possible, d'abord car les items ne sont pas identiques dans les deux évaluations, mais aussi parce que le paradigme méthodologique n'est pas complet avec CEDRE. La DEPP s'appuie sur une évaluation des compétences à partir de « questions scolaires », en d'autres termes des questions correspondant au programme vu en classe par des élèves de 3^e ou de CM2 pour ce qui est de la fin du cycle primaire.

³⁸⁴ Entretien avec une statisticienne, en charge de PISA, DEPP, 1er juillet 2020.

Cette non-récupération de ce qui fait de PISA une évaluation unique en son genre, le choix d'une classe d'âge plutôt que d'un niveau scolaire, est tout à fait importante car elle donne à voir l'objectif mesuré par CEDRE. Du point-de-vue de la définition des résultats attendus de l'éducation, CEDRE reprend l'apprentissage d'un programme comme référence de construction de l'échelle de compétence sur laquelle sont positionnés les résultats des élèves. La performance devient dans CEDRE la maîtrise de compétences définies dans les programmes vus en fin d'un niveau scolaire donné, et plus, comme c'est le cas dans PISA, la capacité à utiliser les apprentissages scolaires dans des situations de vie quotidienne qui ne sont pas vues en classe. Les responsables du bureau de l'évaluation des élèves de la DEPP le disent eux-mêmes clairement lorsqu'ils présentent CEDRE et ses objets d'évaluation dans les rapports techniques des différents cycles. Dans le rapport sur le cycle de mathématiques de 2014, ils indiquent ainsi :

« Ces programmes d'évaluations sont des outils d'observation des acquis des élèves pour le pilotage d'ensemble du système éducatif (Trosseille & Rocher, 2015). Les évaluations de CEDRE (Cycle d'Évaluations Disciplinaires Réalisées sur Échantillons) révèlent ainsi, en référence aux programmes scolaires, les objectifs atteints et ceux qui ne le sont pas. Ces évaluations doivent permettre d'agir au niveau national sur les programmes des disciplines, sur l'organisation des apprentissages, sur les contextes de l'enseignement, sur des populations caractérisées. »³⁸⁵

CEDRE est bien construite comme une évaluation de l'ensemble du système éducatif même si elle se fonde sur l'acquisition des programmes par les élèves. Les usages qui peuvent être faits des résultats sont relatifs pour la DEPP au contenu des programmes d'enseignement d'abord, avant la façon dont ils sont enseignés – très probablement du fait du principe fondamental de liberté pédagogique des enseignants qui limite les instructions en la matière au niveau national – et mettent en lumière des difficultés spécifiques pour certaines

³⁸⁵ DEPP, *Cycle des évaluations disciplinaires réalisées sur échantillons, rapport technique, Mathématiques 2014, Ecole*, décembre 2015, p.3.

catégories d'élèves nécessitant d'en creuser les raisons et les leviers possibles pour les résoudre.

L'évaluation de l'acquisition d'un contenu appris en classe et déterminé au niveau national au sein d'un programme par discipline reste avec CEDRE ce qui permet de déterminer les acquis attendus des élèves de CM2 et de 3^e. En ce sens, la DEPP ne rompt pas avec la tradition d'évaluation française. Ce programme commun à tous les élèves jusqu'à la fin du collège reste en effet central dans la conception française de l'école républicaine post-loi Haby³⁸⁶, qui a contribué à l'unification des filières jusqu'en fin de collège pour aboutir à l'idéal de démocratisation de l'enseignement que fût le collège unique. Objet apolitique³⁸⁷, il porte en lui l'enjeu de fournir à tous les élèves, quelle que soit leur origine sociale et économique, la capacité à acquérir des connaissances similaires au terme de la scolarité obligatoire. Si le principe de son existence n'est pas remis en cause par ceux qui travaillent sur ce référentiel commun, la nécessité d'en faire un outil souple, adaptable aux exigences nouvelles d'une société en mouvement permanent, est régulièrement rappelée³⁸⁸. Dans le cas de CEDRE, l'évolution des programmes nécessite de faire évoluer les items en conservant une comparabilité dans le temps, malgré une évolution du contenu des programmes et des changements de conception des principes directeurs qui les sous-tendent. C'est le cas notamment de l'intégration à partir de 2008 de l'approche par compétence dans la façon de concevoir les items, définie par la loi de 2005 sur le socle commun de connaissances et de compétences comme une combinaison de trois paramètres complémentaires : les connaissances, les capacités et les attitudes³⁸⁹. De nouvelles compétences sont venues s'ajouter, à la fin des années 2000, à celles

³⁸⁶ Loi du 11 juillet 1975.

³⁸⁷ Dominique Raulin, « De nouveaux rapports entre science et politique : le cas des programmes scolaires », *Revue française de pédagogie*, 154, 2006.

³⁸⁸ Voir par exemple Pierre Bourdieu et François Gros, *Principes pour une réflexion sur les contenus de l'enseignement*. Commission présidée par Pierre Bourdieu et François Gros, 1989 Paris, Imprimerie Nationale, 14p. ; Viviane Isambert-Jamati, *Les savoirs scolaires : enjeux sociaux des contenus d'enseignement et de leurs réformes*, 1990, Paris, Ed. Universitaires.

³⁸⁹ Voir notamment la description faite de CEDRE 2009 sur le site de l'éducation nationale : <https://www.education.gouv.fr/cedre-2013-grande-stabilite-des-acquis-en-sciences-en-fin-d-ecole-depuis-2007-8408>.

présentes dans les premiers cycles de CEDRE. Elles visent à évaluer non seulement la capacité à mémoriser des connaissances (« connaître ») et à traiter des documents (« raisonner »), mais aussi à mettre en place une démarche d'investigation, notamment dans les disciplines scientifiques. A partir de 2013, la notion de « tâche complexe » nécessitant à la fois la mobilisation de ressources internes (connaissances acquises en classe, concepts vus) et de ressources externes, comme des documents, est intégrée pour correspondre le mieux possible à la notion de compétence³⁹⁰. Si CEDRE reste fondé sur les programmes scolaires, il ne vise plus à appréhender un ensemble de contenus linéaires et divisés entre disciplines hermétiques et rapproche l'évaluation de la DEPP de la « philosophie » que ses agents attribuent à PISA. L'apparition, y compris dans les disciplines scientifiques, de tâches complexes, est pour certains agents de la DEPP représentative de ce rapprochement. L'un d'entre eux, par ailleurs enseignant en mathématiques, affirme ainsi en entretien :

« C'est vrai qu'au début quand je suis arrivé, on aurait pu ressentir une vision assez critique de ce que PISA pouvait mesurer sur le système éducatif français. Globalement parce que le cadre d'évaluation de PISA est très éloigné de ce qui pouvait être fait en classe du moins en mathématiques. Mais peu à peu, l'évaluation de 2003 a commencé à faire bouger les choses. En 2003 la majeure était mathématiques aussi, on a pu sentir que la réforme des connaissances et compétences de 2005 prenait largement sa source conceptuelle disons, au moins partiellement, dans PISA 2003, notamment en mathématiques. PISA était mentionné dans les textes du programme de la production du socle. Mais évidemment entre la demande institutionnelle et la pratique de classe il y a une très très très longue inertie, et donc le socle n'a pas été réellement mis en œuvre dans les classes avant 2010, même si les décrets datent de 2005.

Le système éducatif français étant souvent piloté par les examens, c'est vraiment la réforme du brevet dans ces années-là, en mathématiques notamment, qui a pu changer les pratiques de classe assez rapidement. Justement sur la notion de tâches complexes. Cette notion de tâche complexe, c'est à dire un problème avec beaucoup de prise d'information, généralement connecté avec le monde réel, où l'élève doit, de sa propre initiative, mettre en relation des connaissances, des savoir-faire,

³⁹⁰ Sandra Andreu, Yann Eteve, Emilie Garcia, Thierry Rocher et Ronan Vourc'h, *Rapport technique CEDRE, sciences expérimentales 2013*, DEPP.

des capacités mathématiques, pour résoudre un problème à question unique, ça, c'est de PISA. Et de plus en plus à partir de 2010 d'ailleurs, des études ont eu lieu sur les énoncés du brevet, il y a une évolution avec une rupture encore une fois dans ces années-là sur les exercices du type PISA, dans le brevet, ce qui n'avait jamais été vu avant. »³⁹¹

On le voit la non-réception supposée de PISA et sa non-influence doit être ici nuancée. Par son influence sur la façon de définir les résultats de l'éducation, PISA a en effet teinté la conception des programmes disciplinaires et les instructions données dans la mise en place des évaluations de fin de cycle, et notamment le brevet pour ce qui est des élèves de 3^e. Par dérivation il semble aussi avoir eu un effet sur la façon de concevoir les items de CEDRE, puisqu'ils ont pour objectif de permettre d'évaluer l'acquisition des compétences prévues dans le programme scolaire.

Les évaluateurs de la DEPP que nous avons rencontrés ont donc montré un intérêt affirmé pour l'évolution du concept de compétence vers une définition plus proche de celle employée par l'OCDE dans PISA. Elle se double d'une validité reconnue dans les faits à la méthode de PISA qui devient très vite, moins de deux ans après le lancement de l'évaluation de l'OCDE, celle de CEDRE. L'expérience méthodologique de PISA vient apporter à la DEPP les compléments techniques qui avaient pu manquer auparavant dans le lancement d'autres études nationales par échantillon, au moment où Claude Thélot était encore directeur de la DEPP. Catherine Moisan, directrice de la DEPP de 2013 à 2015 et par ailleurs agrégée de mathématiques, a notamment insisté en entretien sur les difficultés qu'avait rencontrés la DEPP quelques années avant PISA pour construire des évaluations par échantillon qui soient solides et qui permettent de comparer l'évolution du niveau des élèves d'un cycle à l'autre :

« CEDRE c'est comme PISA, c'est aussi des échantillons avec des disciplines qui tournent. Il y avait avant les évaluations de la DEPP, de Thélot, en CE2 et 6^e qui étaient aussi sur échantillon. Elles avaient quand

³⁹¹ Entretien avec Franck Salles, Chargé d'étude sur l'évaluation des élèves, National Project Manager, DEPP, 11 juillet 2017.

même un gros gros défaut, c'est qu'on était moins costauds, et que c'était peu comparable d'une année sur l'autre. PISA, c'est très sérieux, il ne faut pas lui faire dire ce qu'il ne dit pas mais il n'y a rien à dire sur la méthode. »

392

Le fonctionnement par QCM notamment est repris, alors même qu'il était l'objet de très nombreuses critiques de la part de la DEPP dans le cadre de son utilisation dans PISA. Finalement, il devient la méthode choisie pour l'évaluation "maison". Les arguments donnés par les agents de la DEPP en charge de PISA et de CEDRE pour expliquer ce revirement méthodologique ne font d'ailleurs par mention des oppositions qu'ils avaient pu exprimer dans les premiers travaux d'analyse des résultats de PISA ou dans la note anonyme fournie au Ministre de l'époque et dont les extraits sont cités en début de chapitre. Le directeur de l'époque du bureau des évaluations des élèves expliquait de cette façon le recours au QCM pour CEDRE :

«En 2002, CEDRE avait déjà commencé avec deux disciplines de compétences générales et de langues. Les questions ouvertes étaient source de complexité et très chronophages pour les corriger. Cela demandait de prendre des vacataires, c'était épouvantable. Tout a été transformé en QCM à l'été. »³⁹³

L'argument avancé par Bruno Trosseille est d'abord celui de la simplicité d'exécution et de correction dans un contexte de budget et d'effectifs réduits en interne à la DEPP. La question de la validité scientifique de l'évaluation de l'acquisition de compétences complexes et de la transposition de tâches données en format QCM ³⁹⁴ est finalement évacuée assez rapidement par nos interlocuteurs et, semble-t-il dans leurs propos, par l'ensemble des experts participants à divers titres à la conception des items.

³⁹² Entretien avec Catherine Moisan, Directrice de la DEPP de 2013 à 2015, 10 juin 2020.

³⁹³ Entretien avec Bruno Trosseille, Directeur du bureau des évaluations des élèves, DEPP, 14 avril 2016.

³⁹⁴ Voir par exemple les travaux de René Duval, sur la conversion de registres dans le cas d'évaluations. René Duval, « Quel cognitif retenir en didactique ? », *Recherches en Didactiques des Mathématiques*, 16(3), 349-382. La Pensée Sauvage éditions, Grenoble, 1996.

Pour d'autres de nos interlocuteurs et notamment Isabelle Robin, en poste à la Dgesco au moment de notre entretien mais ancienne de la DEPP au début du lancement de PISA, et qui avait participé aux premières analyses et aux premiers échanges autour de la validité de PISA pour le système français, l'explication est aussi à trouver dans la rassurance des résultats PISA sur la capacité des élèves français à répondre au QCM.

« Parce qu'il faut se souvenir, enfin vous vous ne pouvez pas mais moi je me souviens très bien, qu'avant PISA, avant 2000, on savait que PISA allait arriver, on disait « oh mais on ne sera pas bon à PISA, parce que PISA c'est majoritairement du QCM et nos élèves ne savent pas remplir du QCM et ils ne sont pas du tout formatés pour ça et ils vont se planter ». Sous-entendu, ils vont performer dans les réponses écrites. PISA arrive et on constate exactement le contraire. On constate que les élèves répondent peu, qu'ils cessent de répondre quand ils ne sont pas absolument sûrs, mais plus pervers que ça, on se rend compte qu'ils répondent massivement et massivement juste aux QCM mais qu'en revanche ils sont muets quand on leur demande des restitutions écrites, même courtes. C'est à dire que ce que nous renvoie PISA c'est le contraire de ce qu'on estimait dans un premier temps. Et ça ça a fait choc chez nous, pour nos évaluations. Pour le reste bon, il n'y a pas eu le même choc. Ce que je vais dire ça n'engage que moi mais on a constaté qu'on était dans la moyenne, on s'est rassurés. »

La pratique du QCM était peu répandue dans l'enseignement en France, et les évaluations nationales n'y avaient pas recours pour rendre compte du niveau des élèves, et encore moins tirer des conclusions sur les résultats de l'éducation en général dans le pays. Les spécificités françaises de l'évaluation des acquis des élèves en France portaient certes sur l'objet évalué par les questions posées, l'acquisition des programmes d'enseignement, mais aussi sur la forme de ces questions supposée permettre une expression construite et impliquer une démarche de réflexion adéquate, tout en testant les capacités à rendre compte de cette démarche à l'écrit. Les attendus des évaluateurs nationaux notamment pour PISA étaient donc de voir les élèves français, non habitués au type d'exercice par QCM dans leurs évaluations en classe, peu briller dans leurs réponses. L'inverse se produit dès les premiers cycles PISA. Au lieu de permettre d'expliquer les

mauvais résultats des élèves français à l'évaluation, le format du QCM incite à s'interroger, au-delà de la maîtrise du programme scolaire vu en classe, sur le rapport des élèves français à l'expression d'une réponse. Et finalement, pour obtenir le meilleur taux de réponse possible, et en attendant que les implications des fortes non-réponses observées chez les élèves français aux questions ouvertes, le QCM se voit conférer une place de choix parmi les outils à la disposition des évaluateurs nationaux. Efficace, il est également légitimé par son usage dans PISA, dont la méthodologie devient rapidement internationalement reconnue comme robuste et innovante, notamment du fait de l'appropriation des résultats PISA par plusieurs des pays participants pour fonder des réformes majeures de leur système éducatif, comme en Allemagne.

Si CEDRE s'est appuyé sur les principes méthodologiques de PISA, jugés non seulement robustes mais aussi représentatifs des particularités du système éducatif français, il s'est également appuyé sur les « ratés » de PISA. Les réflexions issues des groupes de travail de PISA ont bénéficié à la réflexion menée en interne dans le processus de construction de la méthode de la nouvelle évaluation « maison ». Ce fût le cas notamment pour l'intégration de variables non-disciplinaires, et la récupération des données concernant la profession des parents des élèves participant à l'évaluation.

« La profession des parents a été dès le départ présente dans CEDRE, avec en plus des variables type motivation, intérêt. On n'a pas fait comme dans PISA, où on avait vu quelques problèmes sur la variable de PISA 2000. La profession du père et de la mère était demandée à l'élève, on lui demandait également de décrire rapidement ce qu'ils faisaient. Ce n'était pas forcément très clair donc on a eu des problèmes de codage. En 2001 on a demandé aux parents leur profession, la corrélation était incertaine. »³⁹⁵

Bruno Trosseille donne ici l'exemple du processus qui a amené la DEPP à peaufiner son indicateur concernant la profession des parents des élèves qui

³⁹⁵ Entretien avec Bruno Trosseille, Directeur du bureau des évaluations des élèves, DEPP, 14 avril 2016.

participent à CEDRE. Le codage identifié comme inexact opéré dans PISA 2000 ne sera pas revu en tant que tel malgré la mobilisation des NPM français auprès de leurs homologues pour faire front commun, comme on l'a vu précédemment. Un indice de statut économique, social et culturel (SESC) sera néanmoins construit pour prendre en compte un ensemble de facteurs, comme le niveau de formation des parents, les biens culturels disponibles à la maison, les ressources pédagogiques à disposition des enfants chez eux, et non simplement la profession des parents déclarée par les élèves. Le travail d'étude approfondie de la façon la plus juste d'obtenir la profession des parents des élèves participants aux évaluations, via des expérimentations nationales, servira à la DEPP pour construire l'indice de position sociale de CEDRE. Elle prendra en compte des facteurs complémentaires à la profession dans l'analyse du poids du facteur familial dans la réussite éducative des élèves. La DEPP s'intéressera ainsi par exemple à l'implication des parents au-delà de leur profession déclarée, pour étudier son effet sur les performances des élèves³⁹⁶. Encore une fois l'évaluation PISA sert de point de comparaison tout au long de l'article présentant les recherches de la DEPP sur le sujet, tout en indiquant bien les limites d'interprétation du questionnaire PISA fondé sur une déclaration des élèves et, pour les auteurs, biaisée du fait de son mode de recueil.

Le rapprochement entre CEDRE et PISA va faire l'objet d'un discours partagé de la part des agents de la DEPP. Ils vont progressivement prendre de la distance par rapport aux craintes et des critiques qu'ils ont pu émettre dans toute la phase préparatoire au premier cycle PISA et au moment de la publication des premiers résultats. La complémentarité de PISA et de CEDRE va être construite et devenir le mot d'ordre général des évaluateurs nationaux à chaque évocation de PISA, dans les documents et publications du ministère comme dans leurs interventions orales lors de colloques ou de forums sur le sujet.

³⁹⁶ Robert Rakocevic, « Implication des parents dans la réussite à l'école : éclairages internationaux », *Education & formations* n°85, novembre 2014.

c) Le discours de complémentarité des évaluations

A la phase de critique méthodologique succède donc assez rapidement un temps de construction d'un discours de complémentarité des résultats de PISA avec ceux de la DEPP, et avec CEDRE en particulier.

Les premières publications de la DEPP sur les résultats du cycle PISA 2000 replacent tout de suite l'évaluation dans l'ensemble des travaux qu'elle mène au niveau national, aussi bien les évaluations bilan en fin de troisième, qui concernent à peu près les élèves de 15 ans, ou les évaluations faites lors de la Journée d'appel et de préparation à la défense ou les évaluations-diagnostic sur lesquelles elle travaille depuis des années. L'intérêt de l'évaluation internationale permet alors, selon les responsables de PISA pour la DEPP, « *d'alimenter le débat sur l'école et des indicateurs pouvant orienter les politiques éducatives* »³⁹⁷. CEDRE n'est pas encore lancé et aucune évaluation interne à la direction ne permet encore d'être présentée comme complémentaire à PISA. Il faut attendre 2006 pour qu'en interne à la DEPP le discours évolue, à la lumière de résultats proches entre CEDRE et PISA. Lors d'un séminaire « les jeudis de la DEPP » organisé en fin d'année 2007, après la publication des résultats de PISA 2006, le directeur de l'époque annonce à ses équipes que les « *résultats rejoignent nos évaluations nationales. Il n'y a aucun biais, aucun reproche ni aucun moyen de se raccrocher à une mauvaise qualité du thermomètre* »³⁹⁸. La similarité des niveaux de compétence des élèves français obtenus à CEDRE et à PISA, alors même que CEDRE a été conçu pour dépasser les critiques du ministère de l'Éducation nationale sur les aspects méthodologiques jugés inadaptés au système français de PISA, oblige les

³⁹⁷ Ginette Bourny, Nicole Braxmeyer, Claire Dupé, Martine Rémond, Isabelle Robin et Thierry Rocher, *Les compétences des élèves français à l'épreuve d'une évaluation internationale, premiers résultats de l'enquête PISA 2000*, DPD, 2001.

³⁹⁸ Dépêche AEF n°88260, citée par Nathalie Mons, Xavier Pons, avec la participation d'Agnès Van Zanten et Jeremy Pouille, *La réception de PISA en France, Connaissances et régulation du système éducatif*, KNOWandPOL, orientation 3, avril 2009.

évaluateurs nationaux comme les décideurs politiques français à renoncer à « casser le thermomètre » comme cela avait pu être fait dans un premier temps.

Ce qui est décrit par certains de nos interlocuteurs comme la « *force de frappe* » de PISA incite à suivre la dynamique induite par la médiatisation croissante du dispositif.

« En arrivant à la DEPP, l'image que j'avais de PISA c'était un palmarès, un classement, un ranking ; et en fait maintenant je ne m'intéresse pas trop à ça mais plus à la manière dont les infos on peut se les approprier, dont on peut créer des indicateurs pour améliorer les politiques éducatives. C'est comme ça que je vois le profit qu'on peut en tirer, et comment ça peut être complémentaire. »³⁹⁹

« PISA effectivement, pour la France, ça a eu beaucoup de répercussions depuis une dizaine d'année sur les grands messages. Le message des inégalités sociales dans l'éducation par exemple, même si on le voit par ailleurs, on n'écoute pas forcément.

PISA ça a ouvert les yeux au politique, c'est vrai qu'il y a un problème. Il en va de l'image de la France auprès de l'OCDE et de la communauté internationale. On doit réagir, mettre en place des dispositifs.

La France a beaucoup pris en compte PISA au détriment d'autres études, nationales, qui ont été menées par la DEPP. C'est bien, ça conforte des résultats. »⁴⁰⁰

Le terme de « *profit* » employé par Ronan Vourc'h est tout à fait intéressant car il donne tout son sens à l'objectif poursuivi par la DEPP. La direction, aux évaluations peu reprises au début des années 2000 dans les moments de réforme budgétaire fondamentale comme la LOLF ou d'évolution du système éducatif français, entend bien utiliser PISA pour retrouver l'influence qu'elle estime être la sienne. Ce nouvel élan vers une reconstruction du discours de la DEPP sur PISA est favorisé par une meilleure réception par l'OCDE des propositions de la DEPP⁴⁰¹, à la fin des années 2000.

³⁹⁹ Entretien avec Ronan Vourc'h, Adjoint au chef de bureau de l'évaluation des élèves, DEPP, 22 juin 2020. .

⁴⁰⁰ Entretien avec Ronan Vourc'h, Adjoint au chef de bureau de l'évaluation des élèves, DEPP, 22 juin 2020.

⁴⁰¹ Xavier Pons, *The French Contribution to the International Design of Pisa*. Know&Pol working note, Paris: OSC, 2009.

« Peu à peu, la France a été de plus en plus écoutée. On a réussi à mettre en œuvre des évaluations de bonne qualité. On fait des contrôles de qualité terribles. On a pu beaucoup plus s'impliquer, on a enfin eu le sentiment d'être entendus. »⁴⁰²

« Progressivement parmi les experts scientifiques il y a eu une plus grande présence française. On a fait des propositions d'experts qui ont été acceptées. On fait passer notre culture, notre façon de voir les choses. On a fait des propositions d'items en science qui ont été acceptées. »⁴⁰³

L'expertise de la DEPP, qu'elle porte sur les propositions de contenu des tests, sur des recommandations d'experts externes (chercheurs ou encore inspecteurs généraux) à impliquer dans les groupes de travail internationaux ou sur les méthodologies employées pour en améliorer la robustesse, leur paraît davantage reconnue par l'OCDE. Cette évolution du rôle des évaluateurs français dans la mise en œuvre de PISA a eu pour effet de rapprocher certaines facettes de l'évaluation internationale de celles d'évaluations de la DEPP, et en particulier de CEDRE. Cela a permis en conséquence d'étudier davantage de comparaisons entre les résultats des élèves français à PISA et les évaluations exhaustives ou bilan, comme CEDRE, proche du point-de-vue de la méthode, voire du contenu puisqu'elle en reprend certains items. Si dans les faits la légitimité méthodologique de PISA semble acquise pour les agents de la DEPP, on perçoit tout de même dans le discours de certains responsables des hésitations quant à la façon de présenter CEDRE par rapport à PISA.

« Ce qui est intéressant c'est que les évaluations se complètent. Il n'y en a pas une à qui on puisse faire tout dire. PISA ça apprend certaines choses sur la France, mais les évaluations exhaustives ça permet aux enseignants d'avoir des actions directes sur le terrain. CEDRE elle permet aussi d'avoir dans différentes disciplines des prises d'info à intervalle régulier, on a des panels qui permettent de suivre les élèves dans le temps. Il y a plusieurs univers. Les enquêtes internationales donnent des infos par rapport à certains prismes, c'est pas les programmes, c'est plus une littératie, la

⁴⁰² Entretien avec Anne-Laure Monnier, Chargée de mission au Conseil supérieur des programmes, 12 septembre 2016.

⁴⁰³ Entretien avec Franck Salles, Chargé d'étude sur l'évaluation des élèves, National Project Manager, DEPP, 11 juillet 2017.

manière dont les élèves s'approprient les connaissances et les mettent en application au quotidien. CEDRE c'est vraiment les programmes. PISA ne va pas nous dire si les élèves français sont au clair par rapport aux programmes. Ça reste une référence le programme, dans CEDRE. Mais bon, nous on fait des évaluations qui visent à évaluer d'autres compétences, la créativité des élèves, des outils sur la manière dont les élèves collaborent entre eux, le numérique, la comm. Il y a une approche plus de compétence, pas forcément liée à un programme, mais transverse. Et aussi l'esprit critique, par rapport aux sources d'information, ça complète l'approche programme qui peut paraître scolaire, réductrice. »⁴⁰⁴

Le constat de base est bien celui d'une complémentarité des travaux, et de l'impossibilité pour un modèle d'évaluation – le modèle de PISA, toujours jugé comme représentatif de la tradition anglo-saxonne - de s'imposer au détriment d'un autre – le modèle français, représenté par la diversité des évaluations qu'ils mènent. Il est donc vain de penser que PISA pourrait supplanter les évaluations de la DEPP, alors que l'évaluation de l'OCDE ne s'intéresse pas au référentiel français des attendus de l'éducation, les programmes d'enseignement, et qu'elle reste par ailleurs déconnectée du monde enseignant. D'un côté il est important de préciser en quoi les évaluations de la DEPP et en particulier CEDRE sont bien différentes de PISA et permettre de fournir une vue plus complète et plus précise sur la situation du système éducatif français et l'évolution des résultats des élèves. D'un autre, il est tout aussi important de montrer en quoi la DEPP a réussi à sortir d'une vision restrictive de l'évaluation de la politique éducative, paradoxalement en s'intéressant à des objets d'études non-disciplinaires, bien différents de l'acquisition de programmes d'enseignement.

La complémentarité entre PISA et CEDRE reste toutefois réduite à ce qui est comparable, et si les agents de la DEPP, par leurs compétences statistiques, maîtrisent l'analyse différenciée des deux évaluations, ce n'est pas nécessairement le cas pour certains observateurs. De la même façon que PISA est trop souvent réduit aux classements des pays qu'il semble fournir sur une échelle linéaire, du meilleur au moins bon, les résultats de CEDRE courent le

⁴⁰⁴ Entretien avec Ronan Vourc'h, Adjoint au chef de bureau de l'évaluation des élèves, DEPP, 22 juin 2020.

risque d'être trop rapidement rapprochés de ceux de son homologue internationale, que ce soit par les décideurs politiques qui s'en empareraient un peu trop rapidement sans en comprendre les fondements statistiques de façon extensive, ou par les médias. Une des statisticiennes de la DEPP impliquée dans le traitement des données des deux évaluations disait ainsi en entretien :

« On ne peut pas faire des croisements parce que c'est pas le même échantillon, mais pour des tendances générales on peut comparer. On a des choses sur PISA avec une vraie chute de scores qu'on ne voit pas chez nous, on a des analyses qui divergent. Mais en général on ne compare pas, et pour une histoire de champs, et parce que la construction n'est pas basée sur les mêmes partis pris. Y compris quand on valorise les résultats on met en garde. Mais on s'en sert quand même. »⁴⁰⁵

La DEPP reprend, en créant cette complémentarité des études, le contrôle de l'analyse des résultats des deux évaluations standardisées. Les deux évaluations ont une base commune, qui facilite la communication autour de CEDRE dans le sillage médiatique de PISA dont il reprend les codes. Néanmoins, les comparaisons sont rendues complexes et techniques du fait de la différence de référentiel de ce qui est entendu par « performance » entre les deux évaluations et du paradigme méthodologique fort du passage à un modèle de réponse à l'item et d'un cadre psychométrique. Les agents de la DEPP, créateurs de CEDRE, en maîtrisent donc la méthode parfaitement, et sont les mieux à même d'en « valoriser » les résultats, d'en tirer des analyses pouvant avoir des implications en termes de pilotage de la politique éducative.

« Les évaluations nationales, c'est une question majeure. Il y a eu un moment où, quand je n'étais pas encore à la DEPP, la Dgesco a fait une erreur énorme. Elle a fait les items d'évaluation. Or elle ne sait pas faire. On ne peut pas piloter et créer ses propres items d'évaluations donc elle ne sait pas faire. (...) Et les évaluations nationales, il y a des dérives sur l'utilisation qu'on en fait. Surtout quand c'est sur échantillon. On a essayé de travailler là-dessus et je pense que ça commence à avoir un effet. »⁴⁰⁶

⁴⁰⁵ Entretien avec une statisticienne, en charge de PISA, DEPP, 1er juillet 2020.

⁴⁰⁶ Entretien avec Catherine Moisan, Directrice de la DEPP de 2013 à 2015, 10 juin 2020. .

Les évaluateurs nationaux font référence auprès du grand public mais aussi auprès des autres directions du ministère de l'Éducation nationale avec lesquelles ils ont parfois été en concurrence pour conserver leur compétence de service de statistique ministérielle, et en particulier la Dgesco.

Il est intéressant de noter que la plus grande implication des agents de la DEPP dans PISA ne s'est pas accompagnée d'une revendication publique de leur rôle plus important dans la conception des cycles de l'évaluation internationale, mais a permis de légitimer le discours de complémentarité entre des études finalement à la main, autant que leur nature le permet, des évaluateurs nationaux. PISA n'est pas présenté comme un travail mené par ou avec l'aide rapprochée de la DEPP. Il est analysé avec distance par les agents de la DEPP comme le travail d'une organisation internationale qu'ils connaissent certes bien et avec laquelle ils sont en contact, mais qui n'est pas le leur. Ce discours distancié, qui permet aux évaluateurs nationaux d'éviter tout lien qui pourrait être fait, par leur auditoire ou leur lectorat, entre leurs évaluations d'une part et les critiques de la méthode de PISA qui continuent d'être entendues dans les médias et certains travaux de recherche d'autre part. L'évaluation PISA, aussi bien étudiée sous l'angle de sa méthodologie que de ses résultats, vient alimenter le travail de la DEPP qui l'utilise de plus en plus souvent pour mettre en lumière ses propres travaux d'évaluation. Lorsqu'elle est utilisée pour éclairer des travaux de fond sur des problématiques jugées saillantes par la DEPP dans le fonctionnement du système éducatif français, elle est d'ailleurs souvent citée au milieu d'autres évaluations internationales, avec pour effet d'en diminuer la prédominance explicative.

Cela évidemment ne concerne pas les notes d'information descriptives qui synthétisent les résultats des élèves français à PISA tous les trois ans. En revanche, cela transparaît dans les productions de la DEPP apportant des éléments statistiques sur la politique éducative, notamment sur la réduction des

inégalités sociales par l'éducation, la mixité sociale dans les établissements scolaires⁴⁰⁷ ou le climat scolaire.

Cela transparaît aussi dans les interventions des responsables de PISA pour la DEPP lors de colloques. L'intervention de Thierry Rocher en février 2018, dans un colloque organisé par le Conseil scientifique de l'éducation nationale intitulé: « Que nous disent les évaluations des élèves en France », reprend l'ensemble des éléments de discours de la DEPP : une connaissance fine de PISA par la DEPP et une proximité avec les experts de l'OCDE, une diminution du pouvoir explication hégémonique de PISA et la réaffirmation de la complémentarité des travaux menés par la DEPP et par les organisations internationales, et enfin la reconnaissance mutuelle de la qualité des travaux menés par la DEPP et par l'OCDE, qui reste un argument d'autorité pour justifier de la qualité technique d'une évaluation statistique.

Thierry Rocher retrouve au banc des intervenants au colloque Eric Charbonnier, analyste à la Direction de l'éducation pour les publications *Regards sur l'éducation* et PISA, et en charge de la communication autour des résultats français. Il le connaît bien, puisque comme il le précise lors de son intervention, « même si les enquêtes internationales sont soit par l'OCDE soit par l'IEA, avec des cahiers des charges qui viennent des institutions internationales, on est très étroitement associé.⁴⁰⁸ » Il est intéressant de noter qu'Eric Charbonnier sera le seul autre intervenant qu'il citera dans son allocution à plusieurs reprises. Par ce biais, il s'agit de montrer que leurs objets d'étude sont proches:

« Comparé aux présentations précédentes, bon à part celle d'Eric Charbonnier où on va être un peu sur les mêmes options, enfin les mêmes types d'activité, je ne vais parler d'expérimentation mais je vais être sur des dispositifs d'observation. »

⁴⁰⁷ Par exemple, l'article publié dans *Education & formations* n°91 sur la massification scolaire et la mixité sociale mentionne en introduction « plusieurs travaux sur le sujet, menés notamment à l'étranger, qui suggèrent la possibilité de décomposer les indices de ségrégation », et la note de bas de page indique que ces travaux sont en réalité ceux menés dans le cadre de PISA, p.54. Pauline Givord, Marine Guillerm, Olivier Monso, Fabrice Murat, « La ségrégation sociale entre les collèges. Quelles différences entre public et privé aux niveaux national, académique et local ? », *Education & formations* n°91, septembre 2016.

⁴⁰⁸ Thierry Rocher, Intervention au colloque du CSEN, « Le rôle de l'expérimentation dans le domaine éducatif » 1er février 2018, Collège de France.

Il s'agit également de légitimer les travaux de la DEPP en rappelant qu'ils sont connus de l'OCDE et utilisés dans leurs analyses sur le fonctionnement du système français.

« Pour les résultats j'ai voulu prendre trois entrées (...). Pour passer à la classe ensuite, avec des points d'alerte sur le primaire, avec deux études récentes d'ailleurs qui ont été mentionnées auparavant par Eric Charbonnier. »

En parallèle de cette proximité affichée avec les experts évaluateurs de l'OCDE, Thierry Rocher explique dès les premières secondes l'importance de considérer un ensemble d'évaluations aux caractéristiques diverses pour conduire l'action publique en éducation. Il insiste sur l'impossibilité selon lui de déduire des liens de causalité d'évaluations « observation », comme peut l'être PISA par exemple, et la complémentarité de ces observations avec les évaluations « expérimentations » qui peuvent être mises en place pour aller plus loin que la simple corrélation de phénomènes.

« Ce serait tout à fait idiot d'opposer expérimentation et observations, qui sont les deux jambes de la démarche qui nous anime pour une meilleure connaissance du système éducatif. »

L'intervention de Thierry Rocher consiste à faire un panorama des évaluations standardisées en France, et il mentionne PISA en le repositionnant dans l'ensemble de toutes les études qui existent, et en montrant que les résultats sont loin d'être complets si l'on s'intéresse par exemple à la problématique saillante de PISA depuis plusieurs cycles, le renforcement des inégalités sociales par l'éducation:

« Les évaluations sur échantillon donnent des informations sur le système dans son ensemble. (...) On a de nombreux dispositifs mis en place. Je les ai divisés en trois catégories. On a tout ce que j'appelle les enquêtes nationales, réalisées sur échantillon, ce sont des enquêtes nationales (...), des enquêtes internationales, on vient de parler de PISA mais il n'y a pas que PISA, il y en a d'autres, (...) et les évaluations nationales exhaustives. »

« On a eu la présentation de PISA qui porte sur les élèves de 15 ans et le concept de littératie. (...) L'OCDE est un des opérateurs d'enquêtes

internationales mais il y a un autre opérateur qui s'appelle l'IEA (...) qui s'occupe aussi depuis très longtemps même des évaluations internationales. »

« PISA c'est à 15 ans. Mais il se passe des choses avant. Si on prend une étude panel, c'est-à-dire un suivi de cohorte, ces élèves quand ils rentrent en 6e ils sont déjà en situation d'inégalité sociale. C'est ce qu'on constate dans toutes nos enquêtes. Et elles s'accroissent si on regarde la progression entre la 6e et la 3e. »

Par cette intervention qui synthétise en moins d'une demi-heure l'ensemble des travaux menés et accompagnés par la DEPP, et les problématiques de politique éducative qu'elles permettent de mettre en lumière, Thierry Rocher déploie sous les yeux des participants au colloque l'ensemble de l'expertise détenu par les agents de la direction, et leur capacité à dépeindre à 360° le fonctionnement du système éducatif et les résultats des élèves.

Ce type d'intervention participe à expliciter et à rendre visible les compétences et les travaux de la DEPP « *encore trop méconnus* » comme le regrette Thierry Rocher en avant-propos de son intervention au colloque du CSEN. Avec la maîtrise technique de ces évaluations, acquise en travaillant sur PISA et rendue légitime dans le champ des évaluations françaises sur les résultats de l'éducation grâce à l'évaluation de l'OCDE, les agents de la DEPP construisent les outils leur permettant de réassurer leur place d'évaluateur unique des résultats de l'éducation en France, dont ils sont les seuls à pouvoir remplir les missions avec toute la rigueur et l'exigence requise.

d) La réaffirmation d'une posture de recherche par les évaluateurs de l'équipe PISA-DEPP par le biais de leur revue à comité de lecture

L'intégration de PISA dans l'ensemble des outils évaluatifs de la DEPP leur permet de valoriser leurs travaux et leur maîtrise technique des méthodes d'évaluations y compris psychométriques auprès non seulement de leurs collègues du ministère de l'Education nationale et des formateurs et enseignants,

mais aussi dans le monde universitaire. La DEPP produit à chaque publication de PISA des notes d'information souvent courtes, de quelques pages, pour reprendre les principaux résultats relatifs aux élèves français. Elles n'ont pas vraiment de dimension analytique — autre que les choix faits par les auteurs de mettre en avant telle ou telle observation dans un document aussi synthétique. Mais au-delà de ces notes, les premiers résultats de PISA ont été contemporains d'une stratégie de valorisation de la revue *Education & formations* éditée depuis plus de 40 ans par la DEPP et dont ses agents rédigent une partie des articles. Cette revue sera le terrain du travail de rapprochement que la DEPP souhaite opérer avec le monde de la recherche. Ce mouvement a deux objectifs. D'une part il vise à légitimer leur maîtrise du fond et de la forme des évaluations nationales et internationales en toute indépendance du pouvoir politique, et donc de PISA au sein de cet ensemble de travaux. D'autre part, il a pour objectif d'améliorer leur visibilité en portant leurs travaux auprès de l'ensemble des acteurs impliqués dans la politique éducative française.

La direction de publication d'*Education & formations* est occupée par la personne à la tête de la DEPP, son directeur ou sa directrice. L'objectif est, selon la DEPP elle-même, de publier « *des analyses et des réflexions (...) propres à faciliter la connaissance et la compréhension du système éducatif français, au fil de son évolution et au prisme de comparaisons internationales* »⁴⁰⁹. Cette publication est destinée à être lue par tous les acteurs du système éducatif, mais aussi par ceux qui travaillent sur les problématiques d'évaluation et qui interviennent dans le champ de la statistique publique. Les sujets sont décidés par la DEPP en fonction des articles soumis par ses équipes ou par des chercheurs extérieurs au ministère qui auraient envoyé des contributions volontaires à la direction en prévision de futurs numéros de la revue. La DEPP, lorsqu'elle choisit par exemple un thème précis, peut solliciter des chercheurs ou des professionnels du champ éducatif afin de leur proposer d'écrire un article. Cette sollicitation est valorisée en son sein et témoigne de l'attention portée aux analyses qui dépassent la superficialité du

⁴⁰⁹ Claire Margaria, Caroline Simonis-Sueur, « La revue fête son centième numéro !: regard rétrospectif sur prêt de 40 ans de publication. », *Education & formations*, DEPP, 2019.

constat qui caractérisent, selon leurs auteurs, les notes d'information à la sortie des résultats PISA. A la manière d'autres services de la statistique ministérielle, comme la DARES pour le ministère du travail et sa revue *Travail et Emploi*⁴¹⁰, la DEPP se rapproche progressivement, avec *Education & Formations*, des codes académiques et formalise la recherche de légitimation de son expertise d'Etat par le monde universitaire⁴¹¹.

L'évaluation PISA y a fait l'objet de plusieurs numéros, dont la régularité est au départ surtout corrélée à la sortie des résultats tous les trois ans, pour devenir plus fréquente à partir du cycle 2012. L'examen des sujets des articles traitant de PISA dans la revue de la DEPP, du statut de leur auteurs (agents de la DEPP ou appartenant au monde universitaire) et leur fréquence au fil des numéros et des ans nous donne à voir la façon dont les évaluateurs de la DEPP, impliqués dans PISA ou en lien avec l'OCDE dans leurs missions, travaillent à construire leur rôle d'interlocuteurs clés pour la recherche sur le sujet des évaluations du système éducatif et de PISA parmi elles, et d'experts capables de solliciter des chercheurs de renom souhaitant apparaître à leurs côtés dans leur publication.

L'évaluation PISA est régulièrement, et de plus en plus régulièrement, étudiée dans les volumes d'*Education & Formations*, que ce soit sous l'angle des analyses qu'elle permet en matière de politique éducative notamment française, ou sous l'angle de la méthodologie qu'elle utilise et de ce qu'elle implique aussi bien du point-de-vue scientifique que du point-de-vue idéologique. Sur la période 2000-2008, soit les trois premiers cycles de PISA, dix articles mentionnant PISA ont été publiés dans la revue. Parmi eux, deux traitaient de la méthode statistique, un s'intéressait spécifiquement aux évaluations nationales et sept utilisaient PISA pour contribuer à une analyse de fond sur des problématiques de résultats de l'éducation au niveau national. La période suivante, de 2011 à 2017, soit les trois

⁴¹⁰ Voir sur cette revue le travail d'Etienne Pénissat et notamment Étienne Pénissat, « Publier des « chiffres officiels » ou les contraintes bureaucratiques et politiques qui façonnent l'expertise d'État. Le cas des statistiques du ministère du Travail », *Gouvernement et action publique*, 4, vol. 1, 2012.

⁴¹¹ Nicolas Belorgey, « Les cadres de l'expertise publique. Les trois dimensions de l'expertise d'État saisies par une ethnographie des économistes de la Banque de France », *Genèses*, 85, 2011.

cycles suivants de PISA, voit ces chiffres fortement augmenter. Ce sont seize articles qui mentionnent PISA, sept portant sur les méthodes statistiques, un sur les évaluations internationales et huit qui invoquent PISA pour apporter des arguments chiffrés aux analyses de fond faites par la DEPP. Si les articles traitaient davantage au début des années 2000 des inconvénients de la méthode de PISA et des innovations qu'elle mobilisait sur le plan conceptuel et sur le plan technique, les résultats ont progressivement été de plus en plus mobilisés par la DEPP pour apporter des fondements statistiques à des études sortant du champ strict de l'état des lieux des compétences des élèves et cherchant à analyser les facteurs pouvant avoir une influence sur les résultats scolaires (le rôle des parents, le climat scolaire, la motivation des élèves, ...).

Chronologie et usages de PISA dans la revue *Education & formations*

La première mention de PISA dans un article de la revue est faite en 2003⁴¹², dans un numéro spécial portant sur des travaux de la DEPP sur les problématiques d'actualité portant sur le système éducatif en général. Les articles sont rédigés conjointement par plusieurs agents de la direction pour apporter des données jugées fiables sur les résultats des élèves et le fonctionnement du système éducatif, en préparation du débat à venir sur l'avenir de l'école. Les chercheurs sont encore à l'époque peu sollicités pour contribuer aux publications, et si la direction anticipe qu'ils liront ses travaux elle prend ses précautions dès le préambule pour se prémunir des possibles critiques sur l'approfondissement limité des analyses faites. L'introduction précise en effet :

« Les rédacteurs ont particulièrement veillé à rendre accessibles les données présentées, sans pour autant remettre en cause leur rigueur scientifique. Certains lecteurs pourront trouver quelques analyses trop complexes, d'autres au contraire les estimeront un peu superficielles. Tout exercice de synthèse nécessite des choix ; les lecteurs qui souhaiteront approfondir un sujet pourront se reporter aux ouvrages

⁴¹² DEPP, *Numéro spécial : 18 questions sur le système éducatif : synthèses de travaux de la DEP*, Education & formations n°66 juillet-décembre 2003.

figurant dans la bibliographie. Bien entendu, les bibliographies ne prétendent, en aucune façon, à l'exhaustivité ; elles ne mentionnent que les principales publications à partir desquelles ont été élaborées ces synthèses. »⁴¹³

Malgré le caractère synthétique des analyses et l'objectif de fournir des éléments de réflexion à un débat national sur l'école souhaitable pour le futur, PISA est tout de même mentionné plusieurs fois, signe de la robustesse reconnue par la DEPP de ses données. PISA est mentionné parmi d'autres évaluations nationales et internationales dans l'article traitant des compétences en lecture des élèves et des jeunes. L'évaluation ne fait pas l'objet d'un traitement en soi et le questionnement tourne toujours autour des compétences disciplinaires. L'évaluation est également mentionnée dans un second article portant sur le collège et interrogeant la réalité du dispositif du collège unique.

Le numéro suivant traitant de PISA, en fin d'année 2004⁴¹⁴, voit apparaître la co-rédaction d'articles sur le sujet par des évaluateurs de la DEPP travaillant avec l'OCDE et par des chercheurs universitaires de renom. Le numéro traite de façon générale de « *questions de méthodes et d'approches novatrices, ainsi qu'à des questions d'équité du système éducatif* ». L'évaluation est mentionnée dans un article traitant du redoublement, corédigé par Thierry Rocher et Olivier Cosnefroy, au titre évocateur : « Le redoublement au cours de la scolarité obligatoire : nouvelles analyses, mêmes constats »⁴¹⁵. Elle vient appuyer une nouvelle fois les conclusions déjà tirées par la DEPP de l'inefficacité du redoublement et est utilisée par les auteurs pour insister sur l'inefficacité, non seulement individuellement pour les élèves, mais aussi pour le système éducatif dans son ensemble. Elle est abordée dans un second chapitre portant sur la méthodologie d'évaluation de la lecture dans les enquêtes internationales, rédigé par Gérard Bonnet⁴¹⁶, chargé de la mission aux

⁴¹³ DEPP, *Numéro spécial : 18 questions sur le système éducatif : synthèses de travaux de la DEP, Education & formations* n°66 juillet-décembre 2003, p.3

⁴¹⁴ *Education & formations* n°70, décembre 2004.

⁴¹⁵ Thierry Rocher et Olivier Cosnefroy, « Le redoublement au cours de la scolarité obligatoire : nouvelles analyses, mêmes constats », *Education & formation* n°70, décembre 2004.

⁴¹⁶ Gérard Bonnet, « Vers une méthode alternative pour l'évaluation de la lecture dans les enquêtes internationales », *Education & formations* n°70, décembre 2004.

relations européennes et internationales de la DEPP, dans lequel il en explique toute la complexité et les apports fondamentaux des travaux d'évaluation nationaux pour sa conception. Elle fait par ailleurs l'objet d'un article à part entière rédigé par un trio de chercheurs, Marie Duru-Bellat, Nathalie Mons et Bruno Suchaut, intitulé « Organisation scolaire et inégalités sociales de performance : les enseignements de l'enquête PISA. »⁴¹⁷

PISA est abordé ensuite en novembre 2008⁴¹⁸, dans un numéro traitant de l'orientation des élèves à la fois du point-de-vue des élèves, des professeurs, et présente des suivis de cohorte à des échelles territoriales différentes. PISA y est utilisé notamment dans un article de Jean-François Mezeix et Catherine Grange du rectorat de Clermont-Ferrand pour éclairer leurs propos sur « Les facteurs influençant l'orientation et le parcours de la troisième au post-baccalauréat d'une cohorte d'élèves de l'académie de Clermont-Ferrand ⁴¹⁹». Les problématiques non-disciplinaires et l'apport de PISA pour éclairer l'orientation des élèves français sont alors placés sous la lumière de la DEPP par cette publication dans la revue officielle de la direction. Un numéro spécial sort par ailleurs le même mois que le précédent et traite cette fois des comparaisons internationales⁴²⁰. Le rédacteur en chef du numéro est Claude Sauvageot, alors chef de la Mission aux relations européennes et internationales à la DEPP, qui travaille en lien étroit avec l'OCDE au sein du projet INES, et qui enseigne par ailleurs à l'époque à l'Université de la Sorbonne et dont les étudiants sont régulièrement orientés vers la direction de l'éducation de l'OCDE pour leurs stages de fin d'étude. La ligne éditoriale choisie est celle du constat de l'omniprésence de plus en plus grande des comparaisons internationales en éducation, témoignant de deux phénomènes

⁴¹⁷ Marie Duru-Bellat, Nathalie Mons et Bruno Suchaut, « Organisation scolaire et inégalités sociales de performance : les enseignements de l'enquête PISA », *Education & formations* n°70, décembre 2004.

⁴¹⁸ *Education & formations* n°77, novembre 2008, 113p.

⁴¹⁹ Jean-François Mezeix et Catherine Grange, « Les facteurs influençant l'orientation et le parcours de la troisième au post-baccalauréat d'une cohorte d'élèves de l'académie de Clermont-Ferrand », *Education & formations* n°77, novembre 2008.

⁴²⁰ *Education & formations* n°78, novembre 2008, 244p.

selon la DEPP : à la fois l'ouverture vers l'étranger et la compétition entre les systèmes. L'objectif du numéro spécial est alors très précisément, dans cette même double logique critique, de « *montrer les progrès réalisés à travers des analyses comparatives, et également de tracer quelques pistes d'amélioration* »⁴²¹. Les publications de l'OCDE sont abordées par la quasi-totalité des articles, qu'il s'agisse de PISA, des données de *Regards sur l'Éducation* ou encore de PIACC sur les compétences des adultes. Nombre d'auteurs ont par ailleurs un lien de proximité avec les travaux de l'OCDE, soit parce qu'ils travaillent en étroite collaboration avec l'organisation, comme Thierry Rocher de la DEPP pour PISA, soit parce qu'ils y ont occupé un poste par le passé, comme Claude Sauvageot ou Norberto Bottani. Ce dernier, présenté comme un consultant indépendant par la DEPP en tête de son article, a été également, pendant plus de vingt ans, responsable des indicateurs des systèmes éducatifs à l'OCDE. Du point-de-vue des analyses permises par les évaluations internationales, PISA est cité par plusieurs auteurs pour étudier les inégalités scolaires ou encore la littératie des adultes. Thierry Rocher y travaille également à nouveau sur la problématique des effets du redoublement sur les résultats des élèves⁴²², étudiée depuis plusieurs années par la DEPP et peu reprise par les décideurs politiques avant que PISA ne formule des conclusions similaires sur l'inefficacité, voire la contre-productivité du redoublement pour les élèves concernés. Un article de Pierre Vrignaud, chercheur à l'Université de Nanterre, étudie la méthodologie de PISA et questionne le pouvoir de la méthode sur la définition des concepts des évaluations menées⁴²³. Il y insiste sur la nécessité de bien expliquer la construction et l'identification des « compétences » mesurées dans PISA, bien souvent sujettes à des raccourcis dans les médias leur conférant une qualité de

⁴²¹ Claude Sauvageot, *Comparaisons internationales*, Education & Formation n°78 numéro spécial, novembre 2008.

⁴²² Thierry Rocher, « Que nous apprennent les évaluations internationales sur le fonctionnement des systèmes éducatifs ? Une illustration avec la question du redoublement », *Education & Formations* n°78, novembre 2008.

⁴²³ Pierre Vrignaud, « La mesure de la littératie dans PISA : la méthodologie est la réponse, mais quelle était la question ? », *Education & Formations* n°78, novembre 2008.

représentation de l'intelligence des élèves, raccourcis pourtant inexacts si l'on regarde la méthode employée.

La mention et l'étude des résultats PISA et de la méthode de l'évaluation internationale s'accélérent à la fin des années 2010. Le volume d'Education & Formation n°79 de décembre 2010 présente un article corédigé par Thierry Rocher et Noémie Le Donné qui traite de la question de la mesure du contexte socio-éducatif des élèves et des écoles et la façon dont elle est questionnée, travaillée par la DEPP à la lumière des apports de la recherche⁴²⁴. La question notamment de l'intégration de la profession des parents des élèves dans la mesure de ce contexte est posée. On retrouve ici les réflexions qui ont fait suite aux tâtonnements du début de PISA et qui ont apportées à l'expertise des agents de la DEPP. L'indice de PISA, présenté comme issu d'un travail de recherche mis en place par Ganzeboom et al. en 1992⁴²⁵, fait l'objet d'une analyse comparative par les auteurs sur sa représentativité, aux côtés d'autres échelles de mesure utilisées dans d'autres évaluations internationales.

Le numéro 80, publié en décembre 2011, voit quant à lui PISA sollicité aussi bien pour sa méthode que pour les analyses qu'il permet au-delà des résultats des élèves à proprement parler. La façon de mesurer l'équité dans PISA est abordée encore une fois par Thierry Rocher et Saskia Keskpaik⁴²⁶. Ils décortiquent l'indice SESC employé par l'OCDE pour représenter les conditions socio-économiques des élèves participant à PISA et concluent à ses limites pour représenter la diversité des situations à la fois des élèves, mais aussi des systèmes éducatifs. Bruno Trosseille, chef du bureau de l'évaluation des élèves de la DEPP et Sylvie Fumel, National Project Manager pour PISA 2009, convoquent quant à eux PISA pour traiter du goût et des habitudes des élèves de 15 ans en matière de lecture.

⁴²⁴ Noémie Le Donné, Thierry Rocher, « Une meilleure mesure du contexte socio-éducatif des élèves et des écoles – Construction d'un indice de position sociale à partir des professions des parents », *Education & Formations* n°79, décembre 2010.

⁴²⁵ Noémie Le Donné, Thierry Rocher, art. cit. p.110. , H.B.G Ganzeboom et al., « A Standard International Socio-Economic Index of Occupational Status », *Social Science Research*, 21, 1992.

⁴²⁶ Saskia Keskpaik et Thierry Rocher, « La mesure de l'équité dans PISA : pour une décomposition des indices statistiques », *Education & Formations* n°80, décembre 2011.

Dans le numéro 85, publié en novembre 2014, Robert Rakocevic de la mission aux relations européennes et internationales de la DEPP étudie avec le lien entre le rôle des parents, et notamment leur implication dans la scolarité de leurs enfants, et les résultats scolaires de ces derniers. Les numéros combinés 86 et 87, publiés en mai 2015, traitent des déterminants exogènes à l'évaluation qui peuvent influencer sur les résultats des élèves. Thierry Rocher et Saskia Keskaik de la DEPP y traitent de « *La motivation des élèves français face à des évaluations à faible enjeu*⁴²⁷ », dont l'évaluation PISA fait partie puisqu'elle n'implique pas individuellement d'évaluation de ses propres résultats pour passer au niveau supérieur. Pascal Bessonneau, Philippe Arzoumanian et Jean-Marc Pastor, également agents de la DEPP, y traitent quant à eux de la comparabilité des résultats d'une évaluation faite sur ordinateur par rapport à une évaluation faite sur papier, en prenant l'exemple de PISA parmi d'autres évaluations qu'ils rangent dans la catégorie de tests « *institutionnels et commerciaux* »⁴²⁸. Les réflexions techniques sur les principes méthodologiques des évaluations des acquis des élèves sont abordées également dans plusieurs articles rédigés par des agents de la DEPP, Bruno Trosseille, Thierry Rocher, Emilie Garcia, Marion Le Cam. La critique de la méthode de PISA n'appartient pas tout à fait au passé puisque les articles concluent que la qualité des indicateurs de la DEPP serait supérieure à ceux notamment de PISA.

Les numéros 88 et 89, publiés en décembre 2015, traitent quant à eux de PISA pour éclairer les débats autour du climat scolaire et de la cyberviolence. Le numéro 90 d'avril 2016, creuse de nouveau la construction des indices de position sociale des élèves par un article de Thierry Rocher encore une fois, en faisant référence aux articles qu'il a pu publier dans d'autres numéros de la revue par le passé. Le numéro 91 de septembre 2016 continue à s'interroger sur

⁴²⁷ Saskia Keskaik et Thierry Rocher, « La motivation des élèves français face à des évaluations à faibles enjeux. Comment la mesurer ? Son impact sur les réponses. » *Education & Formations*, n°86-87, mai 2015.

⁴²⁸ Pascal Bessonneau, Philippe Arzoumanian et Jean-Marc Pastor, « Une évaluation sous format numérique est-elle comparable à une évaluation de type « papier-crayon » ? », *Education & Formations*, n°86-87, mai 2015, p.160.

les déterminants favorisant la reproduction voire l'aggravation des inégalités par l'école, en analysant le phénomène de ségrégation sociale dans les collèges à la lumière notamment de la définition et des données de PISA sur la « *ségrégation sociale entre établissements dans plusieurs pays de l'OCDE* ⁴²⁹ ». Florence Lefresne, agente de la DEPP, y écrit également un article au titre parlant : « *du bon usage des comparaisons internationales dans l'aide au pilotage des systèmes éducatifs* ». Les résultats de PISA y tiennent une place de choix, notamment parce qu'ils auraient selon l'auteur permis « *une prise de conscience collective des inégalités de performances à l'école, et notamment des inégalités sociales de réussite scolaire* ». A quelques semaines de la publication des résultats de PISA 2015 – prévue à la fin de l'année 2016 – le sujet a son importance pour la DEPP qui repositionne encore une fois les travaux de l'OCDE au sein de ses propres travaux. La direction serait chargée de faire le tri entre les informations « pertinentes » issues des évaluations internationales, et celles qui ne le seraient pas ou en tout cas pas à l'instant T de leur sortie :

« La DEPP est fortement impliquée dans la production de données de comparaisons internationales, par son expertise technique et statistique, mais aussi par son engagement actif dans les comités ou réseaux qui les produisent, ainsi que par son investissement dans leur appropriation par les acteurs publics et plus généralement par le grand public. Historiquement, il lui revient d'ailleurs d'avoir sensibilisé la communauté éducative à la lecture des indicateurs internationaux à travers la publication de L'état de l'école, à partir de 1991, alors même que se mettait en place l'ouvrage de référence des indicateurs de l'OCDE que constitue *Regards sur l'éducation*. Depuis elle intègre régulièrement dans ses publications des données comparatives, lorsque ces dernières existent et sont pertinentes. »⁴³⁰

Le numéro 93 de mai 2017 voit apparaître PISA aux côtés de l'évaluation CEDRE pour analyser le lien entre niveaux de littératie et motivation au test. Il est intéressant de noter ici que les déterminants de la réussite éducative

⁴²⁹ Voir la note de bas de page n°3, p.54, dans Pauline Givord, Marine Guillerm, Olivier Monso, Fabrice Murat, « La ségrégation sociale entre les collèges », *Education & Formations* n°91, septembre 2016.

⁴³⁰ Florence Lefresne, « Du bon usage des comparaisons internationales dans l'aide au pilotage des systèmes éducatifs », *Education & Formations* n°91, septembre 2016, p.134.

étudiés par la DEPP dans ses travaux, par exemple la motivation des élèves à répondre à un test standardisé, ne sont plus seulement analysés par le prisme de l'enjeu des collégiens à répondre à PISA, mais plutôt à celui de répondre à CEDRE, dont la méthode est bien explicitement rapprochée de celle de PISA en introduction des propos des auteurs⁴³¹

Les publications de la DEPP témoignent de plusieurs phénomènes. D'abord, par cette revue et la sollicitation de chercheurs et de spécialistes des politiques éducatives, la DEPP se pose en acteur de coordination de la recherche en cours sur les évaluations en éducation, et en particulier les évaluations internationales. Les publications des agents de la DEPP, qu'ils participent à PISA directement comme Thierry Rocher ou qu'ils utilisent les données dans le cadre de leurs évaluations comparées comme Claude Sauvageot, sont intégrées à des dossiers composés d'articles universitaires allant au-delà d'une simple description de résultats d'évaluations des acquis des élèves, se plaçant aussi bien dans le champ de la sociologie de l'éducation, de l'éducation comparée que de l'action publique en éducation ou de la statistique. Les travaux de la DEPP y figurent au même titre que les autres, acquérant une dimension « recherche » mise en avant par les auteurs.

La combinaison de leurs articles avec ceux de chercheurs renommés est pour la DEPP le gage d'une qualité scientifique de ses travaux et le pas vers une prise en compte de leurs évaluations par le monde universitaire. Le centième numéro de la revue mentionne ainsi à l'occasion d'un article spécial le crédit qui

⁴³¹ Sylvie Fumel, Saskia Keskaik, DEPP, « La motivation des élèves à répondre à un test standardisé. Résultats d'une étude dans le cadre de Cedre compétences langagières et littératie », *Education & Formations* n°93, mai 2017. La première phrase de l'article des deux auteurs travaillant à la DEPP est la suivante : « *Les évaluations standardisées des élèves, telles que Cedre ou PISA, renvoient à des enjeux politiques croissants, alors qu'elles restent à faible enjeu pour les élèves y participant.* »

lui est alloué grâce aux contributions des chercheurs et du comité de lecture indépendant :

« Nous espérons ainsi continuer de susciter chez nos lecteurs la curiosité, l'enthousiasme et la satisfaction face aux sujets abordés et à la qualité des articles proposés. Nous souhaitons également remercier tous ceux qui ont fait, et qui font vivre la revue au cours de toutes ces années : les auteurs en premier lieu sans qui la revue ne pourrait pas paraître, mais également les experts, garants de la qualité scientifique des contributions qui nous sont soumises. »⁴³²

Les entretiens que nous avons menés ont confirmé l'envie pour les évaluateurs de la DEPP de créer un pont entre leurs travaux et le monde de la recherche en politique éducative. Plusieurs d'entre eux nous ont en effet proposé franchement de nous partager des indicateurs ou des statistiques brutes afin qu'elles puissent servir de matériau de recherche, et ont insisté sur leur volonté de permettre à leurs articles d'être saisis par les chercheurs pour nourrir leurs analyses.

« Nous on a vraiment à cœur de travailler avec les chercheurs et les doctorants, on a vraiment à cœur à ce que nos données soient partagées. Je voyais vraiment toutes les discussions que j'ai comme celle d'aujourd'hui comme le début de quelque chose. Après il faut voir en fonction de vos besoins ce que vous voudriez avoir en termes de données etc. »⁴³³

Le même message est passé directement aux chercheurs présents au colloque organisé par le Conseil scientifique de l'éducation nationale en février 2018 auquel participe Thierry Rocher :

« J'en profite pour faire de la pub auprès des chercheurs dans la salle. Les panels sont extrêmement riches, c'est longitudinal, il y a des données de compétences, des données sur les familles, des données scolaires etc. J'invite vraiment les chercheurs à considérer qu'ils sont en libre accès pour la recherche. »⁴³⁴

⁴³² Claire Margaria, Caroline Simonis-Sueur, *art. cit.*, p.19.

⁴³³ Entretien avec Franck Salles, Chargé d'étude sur l'évaluation des élèves, National Project Manager, DEPP, 11 juillet 2017.

⁴³⁴ Intervention de Thierry Rocher, colloque organisé par le CSEN, février 2018.

Par cette réorientation de leurs travaux d'analyse, la mise en place en 2014 d'un comité de lecture indépendant, et la distinction claire entre le format de la revue et celui des notes d'information qu'ils peuvent publier par ailleurs, les évaluateurs nationaux renforcent leurs actions pour être distingués de la communication officielle du ministère de l'Éducation nationale. Cela avait été pointé du doigt en 2012 par l'inspection générale de l'INSEE comme un facteur pouvant expliquer sur la baisse de diffusion des publications de la DEPP au début des années 2000, la confusion possible entre les commanditaires ayant pu inciter la DEPP à décaler certaines publications pour éviter qu'elles ne sortent en même temps que des communications politiques sur des sujets proches⁴³⁵.

Conclusion du chapitre

Dès la mise en place des premiers cycles de PISA, la DEPP se positionne en experte du système éducatif français pour alerter sur la faible significativité des résultats de l'évaluation pour la France, du fait de la méthode choisie par l'OCDE. Les critiques de la DEPP portent très spécifiquement sur les aspects de la méthode sur lesquels ils ont eu le moins de prise alors qu'ils étaient en désaccord avec les agents de PISA-OCDE. Elles se centrent en particulier sur les choix de méthode de traduction des items et le recueil de la profession des parents des élèves à partir de laquelle sera construite l'analyse de l'effet de l'éducation sur les inégalités sociales, très médiatisée à partir des années 2010. Par ces critiques, les évaluateurs nationaux revendiquent la maîtrise et la défense d'une tradition d'évaluation française. Les compétences valorisées sont à la fois celles de la DEPP dans son ensemble, comme service statistique ministériel depuis plusieurs décennies, mais aussi celles individuelles des agents qui la composent. C'est bien leur expertise qui leur permet de lancer l'alerte sur les biais de représentativité des résultats de PISA appliqués à la France qu'ils identifient dès le premier cycle,

⁴³⁵ Joël Allain et Michel Devilliers, *Rapport sur la situation des publications de la DEPP*, INSEE, 6 février 2012.

voire en anticipation des premiers tests. PISA débute quelques années seulement après l'échec de IALS, qui avait abouti au retrait de la France de l'évaluation déjà menée par l'OCDE et déjà suivie par la DEPP en interne au ministère de l'Education nationale.

A cela s'ajoutent des critiques générales sur l'approche psychométrique appliquée dans PISA, et associée au modèle d'évaluation anglo-saxon. Ces critiques sont partagées par l'ensemble des acteurs de l'éducation nationale, qu'ils soient agents de la DEPP ou d'autres directions du ministère, enseignants, inspecteurs, formateurs. Le recours aux QCM notamment est jugé, dès les premiers cycles, peu adapté à la philosophie de l'enseignement en France plutôt tourné vers la formulation de réponses sous forme de phrases. Il ne s'agit plus ici d'un affrontement entre individus pour légitimer leurs compétences ou leur expérience de l'évaluation. Le conflit en jeu est celui qui oppose des approches de l'évaluation des élèves qui se sont construites au cours du siècle précédent pour aboutir à des traditions évaluatives bien différentes. Avec son approche psychométrique par QCM, l'OCDE ranime les critiques anciennes formulées à l'égard des travaux d'Alfred Binet à la fin du XIXe siècle et vient remettre au premier plan une méthodologie d'évaluation considérée, par la DEPP et par la grande majorité des acteurs de l'éducation nationale, comme celle des pays anglo-saxons. Alors que la France aurait, grâce à l'expertise de son ministère et de sa direction de l'évaluation, de la performance et de la prospective, développé des méthodes robustes d'évaluation des élèves pouvant bénéficier aux autres systèmes éducatifs des pays participants à PISA, ses alertes nombreuses et soutenues par certains autres pays participants auraient été ignorées. Allant même au-delà de ces critiques techniques, la DEPP se pose en stratège du champ de l'évaluation et suspecte l'OCDE de vouloir supplanter le modèle d'évaluation français y compris dans les pratiques d'évaluation nationales. Loin d'être vus comme des lieux d'échange constructif, comme le souhaitait Andreas Schleicher lorsqu'il narre le mythe créateur de PISA, les espaces de construction de l'évaluation de l'OCDE ont été vécus par les participants français comme des

lieux politiques, d'influence, de démonstration de force, où ils n'ont pas eu la possibilité de valoriser leur expertise comme ils l'auraient souhaité.

Face à ces échecs jugés politiques plus que techniques par les évaluateurs français, ces derniers ne choisissent cependant pas la voie de la dissidence ou de la défection. Parce qu'ils jugent que la France ne peut pas ne pas mettre en œuvre PISA, ils tentent alors de s'approprier le dispositif, à la fois pour ne pas perdre totalement le contrôle, mais aussi pour valoriser leurs propres travaux et impulser de nouvelles évaluations nationales renforçant leur rôle auprès des décideurs politiques. Après le temps de l'évitement, vient donc pour la DEPP le temps de l'usage légitimant, pour apporter un éclairage nouveau sur leurs propres travaux et repenser leurs outils d'évaluation, à la fois sur le fond (objectifs, sujets abordés, résultats attendus) et sur la forme (méthodologie déployée, temporalité des évaluations, constitution des échantillons). Si PISA doit se faire, il faut en tirer profit autant que possible pour la direction. La méthode PISA est donc reprise par la DEPP pour lancer de nouvelles études et s'éloigner de la position d'exécutant de méthodologies d'évaluation pensées par d'autres. C'est le cas de CEDRE, lancé dès les premières années de la décennie 2000. PISA est utilisé par les évaluateurs de la DEPP comme un argument d'autorité méthodologique pour justifier de l'adoption de la démarche psychométrique pour CEDRE, tout en servant de contre-exemple lorsqu'il s'agit de déterminer l'objet principal de l'évaluation de la DEPP qui reste finalement centré autour des programmes d'enseignement par discipline. La DEPP crée avec CEDRE une évaluation qui vise à supplanter la puissance qu'elle perçoit dans PISA, qui soit non seulement innovante et prospective mais aussi adaptée au système éducatif français et donc avec un pouvoir explicatif bien supérieur à son homologue internationale.

CEDRE devient l'outil principal du recentrage que souhaite opérer la DEPP des résultats de PISA sur l'échelon national. Les deux évaluations étant proches, voire même identiques sur certains items, il devient possible de « vérifier » les conclusions de PISA et de ne conserver plus que les conclusions correspondant à celles des évaluations nationales, tout en contrôlant les usages. Les résultats français de PISA, qui peuvent au bout de quelques cycles faire l'objet d'une

analyse temporelle suffisamment riche pour laisser de côté les comparaisons internationales, sont largement utilisés à partir des années 2010 pour alimenter les travaux de fond de la DEPP et venir en éclairage de leurs évaluations nationales. PISA devient un argument de plus dans un panel d'évaluations complémentaires, ce qui en diminue, dans le discours des évaluateurs nationaux en tout cas, la prédominance que peuvent encore lui accorder les médias et les décideurs politiques.

La création d'un ensemble diversifié d'études nationales et internationales, bilan et diagnostic, exhaustives et sur échantillon, ponctuelles et en cohorte, à différents niveaux scolaires, confère à la DEPP la maîtrise d'un volume de données très conséquent sur le système éducatif français. Elle en accentue la communication, notamment auprès des chercheurs et surtout à leurs côtés, dans sa revue *Education & formations* qui se dote d'un comité de lecture au début des années 2010 et accélère son rythme de publication. Par les articles de fond qu'ils publient aux côtés de chercheurs, et qui traitent de plus en plus souvent de PISA comme apport – parmi d'autres évaluations – à leurs analyses des problématiques rencontrées par le système éducatif français, par les interventions qu'ils font lors de colloques scientifiques, ils renforcent la dimension de recherche de leurs travaux qu'ils percevaient chez leurs homologues de l'OCDE sans avoir jusque-là réussi à la développer pour valoriser leur propre expertise.

Partie 3: La stratégie à l'épreuve de l'usage dans la fabrique de la loi et le contrôle de l'action de l'Etat.

Cette partie est consacrée à l'étude de l'évaluation comme outil d'aide à la décision politique : l'effet de PISA sur les relations qu'entretiennent les décideurs politiques avec leurs évaluateurs nationaux, et l'irruption d'évaluations internationales dans la fabrique de la loi nationale, au détriment des experts « maison ». Les orientations de la fabrique de la loi sont construites et données par le gouvernement en place en la personne du ministre de l'Education nationale. C'est dans ce processus, entre l'opération de mise à l'agenda des problèmes publics décrite par Erik Neveu et celle de leur transposition dans un argument de débat de politique éducative, que nous nous plaçons dans cette troisième partie.

Nous regarderons la façon dont les références externes⁴³⁶ que constituent les résultats des élèves français à PISA sont utilisées par les acteurs politiques des politiques publiques, en comparaison de l'usage des références nationales de la DEPP. On ne cherchera pas ici à regarder l'ensemble des usages qui ont pu être faits des conclusions tirées de PISA par l'OCDE ou des conclusions de la DEPP sur ses propres évaluations. Ce qui nous intéresse est de voir en quoi les usages des travaux de la DEPP, la production de son expertise mise au service des décideurs politiques dans le cadre de la fabrique de la loi ou du contrôle des politiques publiques effectués par les parlementaires notamment, donnent à voir la façon dont les décideurs politiques perçoivent la nature de l'expertise des

⁴³⁶ Iris Santos et Jaakko Kauko, « Externalisations in the Portuguese parliament : analysing power struggles and (de-)legitimation with Multiple Streams Approach, *Journal of Education Policy*, juin 2020.

évaluateurs, et par un effet performatif leurs confèrent un rôle qui peut être différent de celui attendu.

Nous ne regarderons pas les raisons qui ont poussé les décideurs à utiliser à certains moments les résultats de PISA et à les laisser de côté à d'autres. Ce sujet a fait l'objet d'études de la part de chercheurs, en particulier sur la première période de PISA, dans la décennie 2000⁴³⁷. Nous interrogerons la mobilisation qui sera faite des productions de la DEPP portant sur les résultats de leurs évaluations, et des résultats de PISA, par ceux qui en sont des destinataires désignés : les acteurs politiques de la politique éducative française.

Alors que la LOLF a définitivement fait entrer l'évaluation dans le cycle de vie des politiques publiques, y compris celle de la politique éducative française, les évaluateurs nationaux qui ont pourtant pour mission « d'évaluer et de mesurer la performance dans les domaines de l'éducation et de la formation »⁴³⁸ sont peu invoqués lorsqu'il s'agit de fabriquer la loi pour réformer en profondeur le système éducatif, contrairement à PISA. Les résultats des évaluations nationales sont peu à peu réservés aux réflexions internes, mais ne sont ni des outils de pilotage privilégiés ni les outils de communication employés aux moments clés de réforme éducative, que ce soit lors des travaux sur la loi de 2005 ou ceux de la loi de 2013. Au contraire, les résultats PISA de la France au cycle 2012 deviennent même l'argument principal qui justifie, pour le gouvernement qui lance la réforme, l'existence même de la loi. De même, alors que les parlementaires voient inscrits dans le marbre de la Constitution leur rôle officiel de contrôle de l'action réformatrice des gouvernements avec la réforme constitutionnelle de 2008, ils ne font qu'assez peu référence aux travaux des évaluateurs nationaux et ne font que rarement appel à eux lors des phases

⁴³⁷ Nathalie Mons et Xavier Pons, avec la participation d'Agnès Van Zanten et Jeremy Pouille, *La réception de PISA en France, Connaissances et régulation du système éducatif*, KNOWandPOL, orientation 3, avril 2009.

⁴³⁸ Article 6 du Décret n° 2014-133 du 17 février 2014 fixant l'organisation de l'administration centrale des ministères de l'éducation nationale et de l'enseignement supérieur et de la recherche.

d'auditions qui jalonnent les parcours législatifs, contrairement aux agents de PISA- OCDE.

Tout comme pour les agents de la DEPP, si la résistance des ministères a été la posture privilégiée au début des cycles PISA - notamment pour s'opposer aux décisions méthodologiques de l'OCDE, pour refuser de compléter l'offre de base de PISA par des indicateurs jugés peu pertinents pour le cadre français (évaluation des compétences en finance par exemple), ou pour résister au verrouillage de la communication autour des résultats assuré par l'OCDE - elle a progressivement cédé la place à une volonté de faire un usage contrôlé des résultats PISA, pourtant peu valorisants, pour justifier de décisions politiques déjà « dans les tuyaux ». Les décideurs politiques utilisent, à la fin de la période, PISA dans la construction de la loi. L'utilisent-ils au même titre qu'auraient théoriquement dû l'être les évaluations nationales? Ou est-elle un outil d'une autre nature ? Pourquoi PISA est-elle l'évaluation référence et pas les résultats de la DEPP qui bien souvent portent sur les mêmes sujets, et sont parfois même publiés avant que PISA ne s'y intéresse ? La DEPP voit alors son influence considérablement réduite. L'une des hypothèses que l'on pose est que l'évaluation PISA, par sa dimension diplomatique et son inscription dans le contexte d'internationalisation des enjeux éducatifs, de réforme de l'appareil d'Etat et de bonne gestion des deniers publics, devient celle à utiliser dans le jeu politique pour montrer aux observateurs internationaux que les enjeux éducatifs français sont compris et sous contrôle.

Nous regarderons, par la double lunette du pouvoir législatif (Chapitre 5) et du pouvoir exécutif (Chapitre 6), en ce que dit l'usage de l'expertise de l'expert qui la produit.

Chapitre 5 : Les parlementaires et les évaluations statistiques de l'action éducative : usage politique des données et expertise de ses producteurs

La DEPP a construit, on l'a vu, une stratégie en plusieurs volets visant à légitimer ses propres évaluations et à replacer PISA dans un mode d'usage plus restreint, tenant compte des limites qu'elle lui attribue pour l'accompagnement de l'action éducative dans le contexte français. Ce cinquième chapitre a pour objectif d'observer l'effet des actions de réaffirmation de l'expertise de la DEPP, au prisme de la façon dont les évaluations nationales et PISA sont utilisées par les parlementaires dans leur mission essentielle de contrôle de l'action du gouvernement et d'évaluation des politiques publiques. Disposant d'un rôle d'élaboration des politiques publiques ainsi que d'une mission de contrôle de l'action du gouvernement, ils participent à la légitimité du système démocratique en assurant le fonctionnement d'un contrôle démocratique permanent⁴³⁹. Pour cela, leur besoin d'aiguillage est constant, et passe certes par les évaluations que le Parlement peut produire et par la connaissance supposée fine des parlementaires des besoins de leur territoire d'élection et de ceux qui y résident. Il passe aussi, et de plus en plus après la réforme constitutionnelle de 2008, par la recherche d'éléments d'analyse pouvant soutenir les alertes qu'ils portent auprès des ministres sur l'application de telle ou telle politique, ou sur les effets à venir de réformes en préparation. En ce sens ils sont – eux et leurs équipes - des

⁴³⁹ Sandra Speer, Valérie Pattyn, Bart De Peuter, « L'importance grandissante des évaluations au sein des parlements : Les gouvernements tenus de rendre des comptes ? » *Revue Internationale des Sciences Administratives*, vol. 81, n°1, 2015.

destinataires privilégiés des évaluations fournissant des données sur l'état et les résultats du système éducatif français dans toutes ses dimensions. De nombreux travaux, français et étrangers, ont étudié la réception des résultats de PISA et sa présence dans les débats politiques nationaux au travers des types d'usage des données de l'évaluation dans les discours politiques et des débats parlementaires⁴⁴⁰, ou de l'apport de PISA dans la rhétorique syndicale des partenaires sociaux majoritaires dans l'enseignement⁴⁴¹. Les données de PISA peuvent être employées à légitimer des réformes, par des raccourcis explicatifs d'ailleurs plus ou moins orientés voire incorrects du point-de-vue de la compréhension des données de l'évaluation. Elles permettent de critiquer, chiffres à l'appui, les réformes entreprises par les gouvernements d'opposition dans le passé. Elles sont utilisées avec plusieurs objectifs pragmatiques dans l'argumentation formulée par les décideurs politiques dans des situations de construction ou de réorientation de la politique éducative. Des recherches menées sur la réception de PISA dans des pays européens voisins de la France ont cherché à expliquer les faibles usages des données statistiques – et donc des évaluations fondées sur ces statistiques - par les parlementaires. L'exigence de transactions politiques constantes serait à l'origine d'une désaffection des acteurs

⁴⁴⁰ Sans fournir une bibliographie complète sur les réceptions politiques internationales de PISA, voici quelques exemples en lien avec les ouvrages et articles utilisés dans cette thèse et qui abordent directement ce questionnement : Xavier Pons, « PISA au Parlement ou l'effet boomerang d'un outil de communication politique », *Administration & Education*, 1, n°145, 2015 ; Natércio Afonso et Estela Costa, *Use and Circulation of OECS's Programme for International Student Assessment (PISA) in Portugal*, Rapport pour le projet *Know&Pol*, Université de Lisbonne, Faculté de psychologie et de sciences de l'éducation, 2009 pour l'utilisation de PISA dans les textes juridiques et les discours des décideurs comme outil de légitimation de mesures proposées au Portugal, ou encore Sotiria Grek, Martin Lawn et Jenny Ozga, *Study on the Use and Circulation of PISA in Scotland*, Rapport pour le projet *Know&Pol*, Université d'Edinburgh, Centre d'études sociologiques, 2009 ; Nathalie Mons et Xavier Pons, « Pourquoi n'y a-t-il pas de choc PISA en France ? Sociologie de la réception d'une enquête internationale dans trois pays européens (2001-2008) », *Revue française de pédagogie*, n°182, janvier-mars 2013 ; C. Morgan, « Tracing the Sub-national Effect of the OECD PISA : Integration into Canada's Decentralized Education System », *Global Social Policy*, 16 :1, 2015.

⁴⁴¹ Voir par exemple Xavier Pons, « Qu'apprend-on vraiment de PISA ? Sociologie de la réception d'une enquête internationale dans trois pays européens (2001-2008) », *Revue internationale d'éducation de Sèvres*, 54, 2010, pp.51-59. ; Eric Mangez, *Réformer les contenus d'enseignement*, Paris, PUF, 2008 ; Nathalie Mons, « Effets théoriques et réels des politiques d'évaluation standardisées », *Revue française de pédagogie*, 169, octobre-novembre-décembre, pp.99-140, 2009.

politiques pour des données rationnelles, parfois trop rationnelles pour permettre les arrangements nécessaires à l'obtention d'un compromis⁴⁴². D'autres travaux ont mis en évidence le rôle des références externes, et parmi elles des résultats des évaluations internationales, dans l'argumentation des parlementaires pour aboutir à travailler de nouveaux sujets de politique éducative au niveau national⁴⁴³. Si beaucoup traitent de PISA comme outil de construction des politiques éducatives nationales lors des débats législatifs au moment de la discussion de projets de loi, peu examinent l'usage des données statistiques et des résultats d'évaluation dans le contrôle quotidien de l'action gouvernementale, pourtant renforcé par les exigences de gouvernance par l'expertise⁴⁴⁴ issues des principes du Nouveau Management Public, et l'obligation de « rendre des comptes » en continu. En cela, l'usage par les parlementaires des évaluations nationales de la DEPP et internationales de l'OCDE avec PISA dans leur veille permanente de l'efficacité de l'action de l'Etat, passée et à venir, nous paraît un objet d'étude intéressant et encore peu étudié en France comme à l'étranger⁴⁴⁵. Par ces différences d'usage apparaît en filigrane le rôle attribué aux différents évaluateurs et la légitimité accordée à leur expertise par ceux qui reçoivent leurs données et recommandations.

Si l'initiative de contrôle de l'action du gouvernement et d'évaluation des politiques publiques était jusqu'alors tacitement reconnue aux parlementaires, par l'intermédiaire notamment des différents moyens d'interpellation en leur possession (questions écrites et orales au gouvernement à échéance régulière, rédaction des rapports et expression à la tribune lors des projets de loi présentés

⁴⁴² Voir par exemple Eric Mangez et Branka Cattonar, « A quoi servent les enquêtes PISA ? Réflexions sur le cas de la Belgique francophone », *Sciences de la société*, n°79, 2011.

⁴⁴³ Iris Santos et Jaakko Kauko, « Externalisations in the Portuguese parliament : analysing power struggles and (de-)legitimation with Multiple Streams Approach », *Journal of Education Policy*, juin 2020.

⁴⁴⁴ Philippe Bezes, Didier Demazière, Thomas Le Bianic, Catherine Paradeise, Romuald Normand, Daniel Benamouzig, Frédéric Pierru et Julia Evetts, « *New Public Management* et professions dans l'État : au-delà des oppositions, quelles recompositions ? », *Sociologie du travail*, Vol. 53 - n° 3, 2011.

⁴⁴⁵ C Pollitt, « Performance information for democracy : The missing link ? » *Evaluation*, 12 (1), pp.38-55 ; Sandra Speer, Valérie Pattyn, Bart De Peuter, « L'importance grandissante des évaluations au sein des parlements : Les gouvernements tenus de rendre des comptes ? » *Revue Internationale des Sciences Administratives*, vol. 81, n°1, 2015.

en séance) et des examens des budgets de l'Etat en chaque fin d'année, elle n'est réellement inscrite dans le marbre qu'au moment de la révision constitutionnelle de 2008⁴⁴⁶, à l'article 24 de la Constitution. Ce pilier de l'Etat de droit⁴⁴⁷ opérationnalise la mission de représentation des citoyens confiée aux parlementaires, en permettant un contrôle total et continu de l'action de l'Etat et des exigences qui lui incombent. Parmi les formes de contrôles à disposition des parlementaires, deux nous intéressent particulièrement dans ce chapitre : les questions au gouvernement et les missions d'information conclues par un rapport d'information. La façon dont PISA et les évaluations de la DEPP sont mobilisées lors des débats en séance, dans le cadre d'un travail de construction législative, sera par ailleurs abordée dans le chapitre 6. Elle se fera par le biais de l'étude de deux grands moments réformateurs du système éducatif en 2005 et 2013 mobilisant les évaluations du système éducatif français.

Les questions au gouvernement sont la façon la plus directe et la plus spontanée d'interroger les ministres sur les résultats de choix de réformes passés, sur des sujets brûlants d'actualité, ou sur des volontés affichées de réformes à venir. Héritées de la III^e République, elles peuvent prendre la forme écrite ou orale. Dans le premier cas elles peuvent être rédigées en quantité voulue par chaque parlementaire, et le gouvernement est tenu d'y répondre dans un délai raisonnable – ce qui n'est en pratique pas toujours le cas, et nombre de questions sont repostées régulièrement par leurs auteurs, notamment en cas de changement de ministre attributaire. Dans le second cas, elles sont posées lors de séances dédiées, directement aux membres du gouvernement présents qui y répondent du tac au tac, sans préparation⁴⁴⁸. Elles obéissent à des règles de temps et de répartition entre les groupes politiques très précises compte-tenu de leur inscription à l'ordre du jour des assemblées et parfois à leur retransmission en

⁴⁴⁶ René Dosière, « Le contrôle ordinaire », *Pouvoirs*, vol 134, n°3, 2010.

⁴⁴⁷ Eric Thiers, « Le contrôle parlementaire et ses limites juridiques : un pouvoir presque sans entraves », *Pouvoirs*, vol 134, n°3, 2010.

⁴⁴⁸ En pratique la courtoisie au sein d'une même famille politique ou lorsque la question est trop importante pour accepter une réponse sans réel fond impliquent que les parlementaires fournissent quelques heures à l'avance leur question à l'équipe du ministre à laquelle elle sera posée.

direct à la télévision. Cette dernière caractéristique leur confère un pouvoir politique particulier, à la fois sur le plan de l'interpellation symbolique et sur celui du choix de sujets d'actualité perçus comme particulièrement problématiques car en nombre très réduit. L'opportunité de poser une question orale est mûrement réfléchi en interne des groupes politiques, et les sujets sont généralement partagés aux parlementaires du groupe en amont pour être validés avant de faire effectivement l'objet d'une question. Les questions au gouvernement, qu'elles soient écrites ou orales, sont généralement rédigées par les parlementaires et leurs équipes de collaborateurs, et ne passent pas entre les mains des administrateurs des assemblées parlementaires, habitués des instances évaluatrices nationales et parfois internationales. Elles ont ceci d'intéressant qu'elles témoignent de la façon dont le politique s'empare des évaluations existantes et donne à voir la place qu'il leur confère dans l'action de contrôle et d'évaluation de l'application des lois et des politiques publiques mises en place par le gouvernement, ainsi que dans le jeu des oppositions politiques.

Deuxième « révélateur » sur lequel nous nous appuyerons dans ce chapitre, les rapports d'information faits au nom des commissions des assemblées parlementaires, qui sont le résultat de missions d'information temporaires menées par les parlementaires sur proposition des Présidents des assemblées. Comme leur nom l'indique, elles ont un rôle théorique d'information mais les sujets qu'elles traitent sont la plupart du temps considérés comme sensibles et d'actualité, intéressant l'ensemble des groupes politiques représentés dans les hémicycles. Les groupes d'opposition ont par ailleurs acquis au début des années 2010 un « droit de tirage » qu'ils peuvent utiliser une fois par session ordinaire pour demander la création d'une mission d'information ou d'une commission d'enquête. Les rapports sont présentés en commission, et peuvent donner lieu à un débat en séance sans vote, ou au dépôt d'une proposition de loi cosignée par les membres de la mission. Certains rapports ont pu même être repris par le gouvernement pour fonder un projet de loi, comme la mission d'information sur les signes religieux à l'école par exemple, menée à l'Assemblée nationale en 2003,

et qui a donné lieu au projet de loi sur la laïcité à l'école⁴⁴⁹. Ces rapports présentent donc l'intérêt d'être lus et débattus de façon transpartisane, portant à connaissance d'un large nombre de parlementaires de toutes familles politiques les évaluations sur lesquels ils s'appuient. Par ailleurs, ils sont archivés soigneusement par les deux assemblées parlementaires et témoignent donc, à leur lecture, de l'évolution de l'influence des évaluations qu'ils utilisent, au fur et à mesure des années.

Nous avons regardé ici l'intégralité des questions posées par les parlementaires de l'Assemblée nationale et du Sénat retranscrites au journal officiel depuis l'année 2000, qui mentionnent l'évaluation Pisa, les travaux de la DEPP⁴⁵⁰ et en particulier CEDRE. Nous avons regardé à la fois les questions écrites et les questions orales qui ont reçu une réponse du gouvernement. Nous avons également parcouru l'ensemble des rapports d'information qui mentionnent PISA et l'évaluation CEDRE, et regardé dans une démarche comparative la façon dont étaient mobilisées les évaluations menées par la DEPP et l'évaluation de l'OCDE.

Nous verrons dans un premier temps que les usages par les parlementaires de PISA et des évaluations de la DEPP contribuent à conférer à l'OCDE et non à la DEPP le statut d'entrepreneur de problème public. Les indicateurs de PISA sont suffisamment nombreux et pratiques d'utilisation pour fonder une argumentation dans une intervention courte pour devenir des outils récurrents lors des séances de question au gouvernement. Leur usage, souvent du ressort de la comparaison internationale sans davantage d'analyse de la signification des positionnements relatifs des pays pour le contexte français, est progressivement assumé par l'ensemble des groupes politiques, de gauche comme de droite, de la majorité gouvernementale comme de l'opposition. Trois années entre 2012 et

⁴⁴⁹ Rapport d'information n°1275 sur la question du port des signes religieux à l'école fait par M. Jean-Louis Debré, Assemblée Nationale, 4 décembre 2003, ayant servi de fondement à la loi n° 2004-228 du 15 mars 2004 encadrant, en application du principe de laïcité, le port de signes ou de tenues manifestant une appartenance religieuse dans les écoles, collèges et lycées publics.

⁴⁵⁰ Nous avons regardé les questions mentionnant selon les périodes la DPD, la DEP, la DEPP, sous forme d'acronyme ou non.

2014 verront néanmoins l'évaluation CEDRE posée sur un même pied que PISA, témoignant d'une réussite même relative de la stratégie de la DEPP d'inscrire PISA dans son panel d'évaluation, toutefois très dépendante des ministres en place.

Nous verrons ensuite comment les usages des outils d'évaluation tels que PISA et les travaux de la DEPP se trouvent orientés par le contexte grandissant d'internationalisation des enjeux de politique publique, de l'éducation à la situation de l'emploi. Cela se fait au détriment des évaluations nationales qui ne permettent pas, par leur nature et les résultats qu'elles évaluent, aux décideurs politiques français de rendre compte de la place de la France dans la concurrence internationale. Malgré les efforts déployés par la DEPP pour faire entendre la valeur de la comparaison nationale temporelle plus que celle d'une lecture simple les classements représentés dans les publications de l'OCDE, le benchmark s'ancre comme une pratique classique des parlementaires lorsqu'il s'agit d'évaluer les résultats du système éducatif français et de tirer des conclusions sur les réformes à mener.

Enfin, nous verrons que l'usage comparé des travaux de la DEPP et de PISA montre en toile de fond la force croissante du débat sur l'indépendance d'une direction d'évaluation interne au ministère de l'Education nationale et donc sur la valeur des résultats des évaluations qu'elle conduit pour étayer les décisions de réforme du système éducatif. Alors que la DEPP ne parvient pas à construire une identité d'évaluateur indépendant expert des politiques éducatives qui puisse convaincre l'ensemble des destinataires politiques de ses travaux, et en particulier les parlementaires, l'OCDE et son PISA devient l'expert le plus légitime pour fournir des données jugées fiables par l'ensemble des groupes politiques.

1) Quelles évaluations pour mettre à l'agenda les problèmes publics ?

La lecture des questions posées par les parlementaires au gouvernement montre des usages différenciés de PISA et des travaux de la DEPP, qui témoignent du rôle que les parlementaires comme les ministres successifs leur assignent. Nous avons commencé l'étude de ces questions au début des années 2000, au moment du lancement de PISA et au moment de l'institutionnalisation de l'évaluation par indicateur dans le sillage de la LOLF. Les dernières questions étudiées datent de la fin de l'année 2019, c'est-à-dire au moment de la publication des résultats de PISA 2018, permettant ainsi de conclure sur la fin d'un cycle d'évaluation. Cet examen des interventions parlementaires et des réponses du gouvernement permettent de tirer trois types de conclusions sur le rôle des évaluateurs dans la construction de l'action publique en éducation.

Cela nous indique d'abord en quoi les résultats sont connus des parlementaires et des ministres, et donc en quoi les évaluateurs ont réussi le passage de la production et de l'analyse des évaluations à celle de leur popularisation auprès de ceux qui contrôlent l'action de l'Etat. Afin d'étudier la façon dont la stratégie de positionnement de la DEPP a ou non eu un effet sur les usages de ses travaux et de PISA par les parlementaires, nous nous intéressons ici au résultat du processus de popularisation, et non au cheminement qui aboutit ou non à porter à la connaissance des parlementaires les résultats de PISA d'une part et des travaux de la DEPP d'autre part.

Cela nous indique ensuite en quoi les résultats des évaluations françaises par rapport à ceux de l'évaluation de l'OCDE sont considérés, par ceux qui évaluent les résultats de l'action éducative, comme suffisamment indépendants et suffisamment représentatifs d'une situation d'urgence pour nécessiter des réponses politiques claires du gouvernement. D'autre part, cela montre en quoi ils sont suffisamment indépendants et représentatifs pour servir de fondement à un argumentaire politique et pour être en particulier, dans le cas des questions orales, sollicités dans le cadre d'intervention en séance publique au Parlement,

auprès de la représentation nationale et devant, la plupart du temps, une partie non négligeable du gouvernement.

Cela nous indique enfin, en regardant l'évolution des usages au Parlement au fil des deux décennies 2000 et 2010, dans quelle mesure les stratégies d'influence de la DEPP et de l'OCDE avec PISA ont pu contribuer à entériner leur rôle de soutien à l'évaluation continue de la politique éducative française.

a) La mise à l'agenda de l'urgent : puissance de l'oral et temps long de l'écrit

L'examen que nous avons mené a porté à la fois sur la forme des questions et sur le fond, et a été guidé d'abord par l'hypothèse d'un rôle induit différent selon que les évaluations sont sollicitées dans le cadre des questions écrites ou des questions orales. On l'a dit en introduction de ce chapitre, les questions écrites peuvent être posées en grand nombre⁴⁵¹ et tout au long de l'année. Les réponses faites par le gouvernement le sont dans des délais très variables. En moyenne, les deux assemblées parlementaires indiquent un délai d'environ 180 jours, qui varie selon les années, pour qu'une question trouve réponse. Si en théorie les questions écrites devaient au début des années 2000 recevoir une réponse sous un délai d'un mois, passé progressivement règlementairement à deux mois au milieu des années 2010, en pratique les réponses sont parfois données plusieurs années plus tard, par des ministres différents, répondant par un même texte à plusieurs questions traitant de sujets approchants. Ce constat a même fait l'objet de

⁴⁵¹ Dans le cas de l'Assemblée Nationale, le dispositif de questions écrites a été prévu en 1909 lors de la révision du règlement de la Chambre des députés pour leur permettre d'interroger le gouvernement sans que cela n'implique une mise en cause publique de sa responsabilité. Prérogative de chaque député, elles pouvaient jusqu'à la révision de l'article 135 du règlement du 28 novembre 2014, être posées en nombre illimité. Désormais, le nombre de questions est plafonné, en contrepartie d'une réponse imposée aux ministres sous deux mois. En début de chaque session ordinaire, la Conférence des présidents fixe le nombre maximal de questions écrites pouvant être posées pendant l'année. Cette révision du règlement fait régulièrement débat au sein des parlementaires car elle est perçue par certains comme une limitation du pouvoir de contrôle qui leur revient. Une proposition de résolution du député Olivier Dassault a d'ailleurs été déposée en septembre 2020 pour revenir sur le plafonnement des questions écrites. Dans le cas du Sénat, le nombre de questions écrites n'est pas limité et le délai de réponse des ministres est de deux mois.

questions au gouvernement demandant aux ministres de bien vouloir répondre dans des délais acceptables aux sollicitations des parlementaires⁴⁵². L'enjeu de la question écrite n'est donc pas le même que celui de la question orale, qui est, elle, posée directement et devant tous les parlementaires présents au ministre le mieux à même d'y répondre. La popularisation du sujet à grande échelle est immédiate, la réponse reçue est immédiate, et la possibilité de communiquer sur l'interpellation du gouvernement sur un sujet jugé problématique et appelant une réponse urgente est immédiate.

Si l'on regarde les questions posées mentionnant la DEPP, ce sont en très grande majorité des questions écrites. Au Sénat, 90% des questions-réponses qui se réfèrent aux travaux de la DEPP sont écrites. A l'Assemblée Nationale, la proportion est même encore plus importante, puisqu'une seule question se référant à la DEPP a été posée à l'oral et a reçu une réponse⁴⁵³. Le temps accordé au gouvernement pour répondre est plus long, sans que l'auteur de la question ne sache bien à quelle échéance il pourra obtenir cette réponse. La question posée signifie bien qu'est identifié un sujet de politique éducative, mais que les parlementaires ne souhaitent pas forcément en faire un sujet public, et le positionner en haut de la pile des priorités du moment. Les modes de questionnement de référant à PISA se répartissent quant à eux bien différemment. Au Sénat, 29 des 40 questions mentionnant PISA sont des questions orales, soit près des trois quarts. A l'Assemblée Nationale, le ratio est plus faible compte-tenu du nombre bien plus important de questions écrites posées, mais ce sont tout de même près d'une question sur cinq évoquant PISA qui l'ont été à l'oral, soit 33 sur les 179 identifiées.

Ces questions orales ont été posées pour une partie d'entre-elles au moment de la sortie des publications donnant les résultats d'ensemble de chaque cycle, tous les trois ans en fin d'année, et pour une autre partie au cours des ans.

⁴⁵² Voir par exemple la question écrite n°6388 de M. Eric Pauget, député Les Républicains, publiée au JO de l'Assemblée Nationale le 13 mars 2018 ; ou la question écrite n°19870 de M. Philippe Paul, sénateur Les Républicains du Finistère, publiée au JO du Sénat le 31 décembre 2020.

⁴⁵³ Question orale n°613 de Madame Cécile Gallez, Nord, UMP, publiée au JO de l'Assemblée Nationale le 14 mai 2003.

La réactivité parlementaire à la sortie de PISA a été en tout état de cause bien plus forte que lors de la sortie – assez confidentielle il faut le dire – des évaluations de la DEPP. Les publications PISA ont donné lieu quasi-systématiquement à des séries de questions écrites et orales interrogeant le ministre sur sa réaction suite aux résultats obtenus par « la France », ou, de façon moins neutre, sur sa réaction aux résultats « traumatisants ⁴⁵⁴», « navrants ⁴⁵⁵», « insatisfaisants ⁴⁵⁶» donnés par PISA. L'observation des questions posées sur PISA au fil des cycles montre une montée en intérêt de la part des parlementaires pour interroger le ministre à la fois sur les résultats de PISA et sur des enjeux précis d'évolution du système éducatif français. Les cycles de PISA 2003 et 2006 suscitent d'abord quelques questions isolées des parlementaires à la fois au Sénat et à l'Assemblée Nationale, à la sortie des publications de l'OCDE. Les cycles 2009, 2012 et 2015 donnent lieu à des groupements coordonnés de questions de plusieurs parlementaires sur le sujet. Elles abordent des indicateurs PISA différents pour interroger le ministre à la fois globalement sur sa réaction aux résultats des élèves français, et plus spécifiquement sur des conclusions ponctuelles de l'OCDE pouvant amener à revoir le fonctionnement du système éducatif français sur le nombre d'enfants par classe, le nombre d'enseignants, la mixité dans les établissements scolaires, l'évolution des méthodes d'apprentissage de la lecture etc. Une séance de questions cribles thématiques⁴⁵⁷ est même organisée au Sénat spécifiquement sur ce sujet le 9 février 2011, permettant à l'ensemble des groupes politiques d'interroger le gouvernement sur les résultats des cycles précédents de PISA.

La publication des évaluations de la DEPP ne donne quant à elle pas lieu à ce type de réactions organisées, même dans le cas de CEDRE qui se rapproche

⁴⁵⁴ Question au gouvernement n°1399 de Monsieur Jean-Louis Borloo, Nord, Union des Démocrates Indépendants, publiée au JO de l'Assemblée Nationale le 4 décembre 2013.

⁴⁵⁵ Question au gouvernement n°1069G de Monsieur Gérard Longuet, Meuse, Les Républicains, publiée au JO du Sénat le 16 décembre 2016.

⁴⁵⁶ Question au gouvernement n°1109G de Monsieur Alain Marc, Aveyron, Les Républicains, publiée au JO du Sénat le 18 janvier 2017.

⁴⁵⁷ Les questions cribles thématiques ont été mises en place d'octobre 2009 à septembre 2015 au Sénat. Deux fois par mois, pendant 45 minutes, neuf questions pouvaient être posées sur un thème unique déterminé par la Conférence des présidents du Sénat. Les thèmes ont été variés, de la réforme de la dotation globale de fonctionnement à la réforme du collège, en passant par la forêt française ou les services à la personne.

pourtant de la forme de l'évaluation de l'OCDE, et qui a été plusieurs fois portée à la connaissance des parlementaires puisqu'on la retrouve citée par plusieurs ministres en séance, notamment au moment de l'examen de plusieurs budgets annuels de l'éducation nationale ou des grandes lois de réforme du système éducatif. L'évaluation n'a été citée que dans une question orale au Sénat depuis 2000, et pour appuyer l'analyse faite des résultats des élèves français à PISA en mathématiques⁴⁵⁸. Elle n'a pas été davantage sollicitée dans les questions orales à l'Assemblée Nationale, n'apparaissant seulement que dans trois questions-réponses en 2011 et 2014.

L'intensité de sollicitation de telle ou telle évaluation lors de questions orales ou écrites n'est pas neutre lorsque l'on veut expliquer l'influence des évaluateurs dans le contrôle de l'action du gouvernement. Un collaborateur parlementaire de la Commission de la culture et de l'éducation du Sénat nous précisait en entretien le mécanisme de choix entre la forme écrite et orale.

« C'est sûr que les questions orales, ça passe à la télé, on peut avoir une réponse tout de suite, souvent un peu de langue de bois quand même, on peut publier sur les réseaux et on peut communiquer. Là, sur PISA, c'est l'occasion rêvée de demander des réponses sur certains points qu'on pourra travailler ensuite dans d'autres projets de loi qui arriveront dans l'année. Donc il faut demander au groupe si on peut avoir une question sur le sujet, généralement ça passe. Sur les études de la DEPP, c'est souvent plus descriptif, moins problématisé, il faut expliquer à tout le monde ce que les chiffres impliquent alors que PISA tout le monde connaît. En deux minutes, c'est le temps de la question qu'on a en gros, c'est plus percutant. »⁴⁵⁹

On retrouve dans cette illustration le rôle assigné à chaque type de question, et par dérivation à chaque évaluation. Le choix de l'écrit ou de l'oral est fait selon les enjeux du sujet abordé, qu'ils soient politiques ou qu'ils relèvent de l'évaluation de la politique éducative menée par le gouvernement, et selon la facilité de compréhension par l'auditoire des données produites par les évaluateurs. A ce jeu, PISA semble, au vu des usages, avoir pris la place de

⁴⁵⁸ Question au gouvernement n°1069G de Monsieur Gérard Longuet, Meuse, Les Républicains, publiée au JO du Sénat le 16 décembre 2016.

⁴⁵⁹ Entretien avec un collaborateur parlementaire du Sénat, groupe de gauche, Commission éducation et affaires culturelles, 22 juin 2020.

référence privilégiée par les parlementaires pour interpellier le gouvernement et communiquer par la suite auprès des citoyens. Les fréquences d'usage des travaux de la DEPP et de PISA viennent appuyer cette conclusion car la différence est flagrante. Si l'on compare strictement évaluation à évaluation, et donc PISA à CEDRE, l'usage de PISA est bien supérieur à celui de sa concurrente nationale. Quarante questions traitent de PISA au Sénat entre 2000 et 2019, contre deux seulement pour CEDRE, toutes formes confondues. À l'Assemblée Nationale, 186 questions traitent de l'évaluation de l'OCDE et 34 de celle de la DEPP. La DEPP est citée plus souvent que PISA, 81 fois au Sénat et 239 fois à l'Assemblée Nationale, mais le type de travaux utilisés est très variable, allant de données chiffrées descriptives sur l'état de l'école aux résultats d'évaluations-bilan et diagnostics des élèves. Quelques questions interrogent le bienfondé de méthodes employées par la DEPP, lors du lancement de nouvelles études par la direction du ministère portant sur des thématiques nouvelles, parfois extra-disciplinaires (l'accès aux produits culturels au domicile des élèves, les activités syndicales des enseignants, ...). Il est intéressant de noter que les parlementaires n'interrogent jamais le bienfondé de la participation de la France à l'évaluation PISA, ni la méthodologie employée, ni encore la supposée origine anglo-saxonne du programme. Ces arguments ne sont par ailleurs jamais, à une exception près, employés par les ministres de tous bords politiques pour expliquer à la représentation nationale les résultats des élèves français, jugés dans le meilleur des cas moyens.

La connaissance même lointaine qu'ont les acteurs politiques de PISA, à la différence des nombreux travaux menés par la DEPP, en font un outil à la fois facile d'accès pour les rédacteurs des questions – des synthèses ou des morceaux choisis sont en permanence sortis dans les médias ou dans les articles de vulgarisation, y compris les notes d'information de la DEPP – et facile à comprendre pour celui qui écoute la question posée. En ce sens, l'OCDE a réussi avec son PISA à développer un outil efficace de promotion de problèmes de politique éducative qu'elle juge nécessaire de porter à connaissance des décideurs politiques. Erik Neveu soulignait dans son analyse des différentes opérations

aboutissant à la constitution d'un problème public : « *Promouvoir efficacement un fait social au statut de problème public, c'est savoir intéresser.* »⁴⁶⁰ Alors que la DEPP semble avoir, y compris à la fin des années 2010, des difficultés à faire entendre sa voix et à valoriser l'intérêt de ses travaux pour le contrôle continu de l'action publique en éducation, PISA transparaît à travers les questions-réponses parlementaires comme étant un sujet d'intérêt de premier ordre dont on ne remet pas en cause ni la méthode, ni les données, ni les conclusions.

b) Effets argumentatifs et mise à l'agenda des problèmes publics

Le second axe de notre analyse des usages comparés des évaluations de la DEPP et de l'OCDE est celui de l'effet argumentatif qu'elles apportent lorsqu'elles sont employées dans les questions. En plus de l'intérêt manifeste des parlementaires pour PISA au sein de la panoplie d'évaluations à leur disposition, l'évaluation de l'OCDE diffère des travaux de la DEPP par les possibilités argumentatives qu'elle permet. L'observation des effets argumentatifs différenciés permet de poursuivre l'analyse des caractéristiques conférant à PISA des conditions d'usages plus favorables dans un cadre parlementaire.

Les questions ont été réparties en deux grandes catégories d'effet: l'effet de justification et l'effet que nous qualifions d'argument d'autorité. Le premier correspond à la catégorie de questions sollicitant PISA ou les travaux de la DEPP comme simple argument pour prouver une critique formulée par l'orateur de la politique éducative menée par le gouvernement, ou pour prouver l'urgence d'agir sur une problématique identifiée par le parlementaire. Il implique de se poser les questions suivantes : L'évaluation employée vient-elle justifier l'argument donné par le parlementaire ? En d'autres termes, nous regardons ici si le parlementaire commence par interroger le ministre sur un problème qu'il perçoit dans le fonctionnement ou les résultats du système éducatif français, et utilise les travaux de PISA, ou de la DEPP pour servir d'argument statistique venant prouver son

⁴⁶⁰ Erik Neveu, *Sociologie politique des problèmes publics*, Paris, Armand Colin, 2015.

argument. Le second effet correspond à la catégorie de questions interrogeant le ministre directement sur les conclusions des évaluations, de PISA d'une part et des travaux de la DEPP d'autre part. La question est alors : L'évaluation est-elle directement l'objet central de la question, interrogeant sur ce qu'elle implique de la part du gouvernement en matière de nouvelles orientations souhaitables de politique éducative ? Elle est alors citée en début d'intervention, et toute la suite de la démonstration du parlementaire se réfère à l'évaluation et n'a de sens que par rapport à elle.

L'examen des questions à la recherche de ces deux effets argumentatifs nous permet de voir que les évaluations de la DEPP et PISA ne sont pas employées à de mêmes fins. Au Sénat, 22 questions faisant référence à PISA utilisent l'évaluation pour justifier un problème pré-identifié, 8 portent spécifiquement sur ce qu'impliquent les résultats PISA en termes de réforme de la politique éducative française. A l'Assemblée Nationale, ce sont 70 questions qui utilisent l'effet de justification de PISA, et 47 qui en font un argument d'autorité. Si l'on regarde les usages des évaluations de la DEPP, on constate que la quasi-intégralité des références sont faites avec un effet de justification. Dans tous les cas, les références aux travaux de la DEPP ont pris la forme essentiellement de données descriptives, chiffrées. Il est finalement fait peu appel à la démarche d'évaluation et de prospective qu'elle porte pourtant dans son nom. La direction devient la pourvoyeuse de chiffres donnant à voir l'organisation et le fonctionnement du système scolaire, dont les parlementaires tirent des conclusions qu'ils jugent alarmantes et qui méritent d'interroger le ministre sur la stratégie prévue. La DEPP montre que la scolarisation des moins de deux ans baisse, que l'absentéisme à l'école augmente, que les enseignants sont remplacés dans des délais longs, que le harcèlement scolaire est à un niveau donné selon les déclarations des élèves, que tant d'heures supplémentaires sont consommées au cours d'une année scolaire donnée, que tant de places sont ouvertes au concours d'enseignant en Education Physique et Sportive... Tous ces chiffres font ensuite l'objet d'une analyse par les parlementaires au prisme de leur couleur politique. Les places au concours d'enseignant en EPS par exemple, sujet de très

nombreuses questions au Sénat comme à l'Assemblée en 2006, sont mises en regard de l'évolution du nombre d'élèves à venir lors de la prochaine rentrée et les conclusions alertent les parlementaires quant au décalage qu'il semble y avoir entre les besoins de recrutement et les places ouvertes au concours d'enseignant. En revanche, peu de références aux travaux de la DEPP donnent lieu à des interrogations sur les orientations futures de la politique éducative du ministre. La dimension prospective des évaluations employées par les parlementaires se borne la plupart du temps à des questionnements sur l'évolution des effectifs d'élèves à venir, ou sur les questions de financement du système. Très peu de questions interrogent le ministre sur les orientations à prendre suite à un problème public « révélé » par la DEPP, comme c'est en revanche régulièrement le cas pour PISA.

L'absence d'attribution à la DEPP d'un rôle d'évaluateur fournissant clé en main aux décideurs politiques de statistiques analytiques traduisibles en politiques éducatives est renforcée par la difficulté qu'a la direction à légitimer ses éventuelles recommandations. Elle en formule relativement peu dans les notes d'information qu'elle produit. La direction tend davantage à orienter l'action publique en sélectionnant les points d'intérêt qu'elle perçoit à partir des résultats de ses travaux statistiques et en en faisant des notes synthétiques spécifiques, qu'en publiant à la manière de PISA des rapports exhaustifs organisés sous forme d'indicateurs problématisés, assortis de conclusions opérationnelles. Cette identité de direction aux missions de production de données descriptives d'un état des lieux est d'ailleurs énoncée clairement par certains parlementaires. En entretien, Yves Durand, député de la commission des affaires culturelles et éducation, estimait même que la dénomination de la direction était trompeuse sur la nature de ses missions, sans pour autant en nier l'importance.

« La DEPP fait de la statistique. Contrairement à ce que son nom indique, elle ne fait pas le travail d'évaluation mais de compilation de statistiques absolument nécessaires.

Je dis pas que ça ne sert à rien, c'est absolument nécessaire pour justement faire l'évaluation. Mais la DEPP en soi ne fait pas

d'évaluation. Quand vous regardez les travaux de la DEPP, c'est l'état de l'école qui est remarquable et sur lequel il faut se baser pour justement faire l'évaluation et des préconisations. Mais ça n'est pas un travail d'évaluation. »⁴⁶¹

Les travaux de la DEPP, et CEDRE ne fait pas vraiment exception, ne fournissent pas de conclusions immédiatement utilisables par les parlementaires dans le contrôle de l'action publique et éducation. L'absence de passage des chiffres à des recommandations de politique éducative, davantage du ressort théorique d'autres directions du ministère et en particulier de la Dgesco, complexifie la mise à l'agenda des problèmes que la DEPP souhaite mettre en avant à l'issue de ses évaluations. Les évaluations qu'elle produit donnent certes à voir, par des statistiques les éventuels défaillances du système scolaire par rapport à ce qui est attendu de lui et inscrit dans la loi. Il est malgré cela difficile pour la direction de positionner les conclusions de ses travaux en haut de la liste des priorités des acteurs politiques, et a fortiori d'être à l'origine de prises de décision en termes d'action éducative⁴⁶².

PISA en revanche est utilisé de deux façons complémentaires au sein des assemblées parlementaires. Il est à la fois un objet de question, et sert également de fondement argumentatif pour interroger sur des questions de fond de réforme du système éducatif, dans une démarche affichée bien plus prospective que les travaux de la DEPP. Il est objet de question d'abord en ce que ses conclusions sont le sujet, au sens grammatical du terme, de nombreuses questions posées par

⁴⁶¹ Entretien avec Yves Durand, Sénateur socialiste puis LREM de Mayenne, rapporteur de la loi refondation de l'école, président du comité de suivi et membre du comité d'évaluation des politiques publique, 14 juin 2017.

⁴⁶² Voir la définition donnée par John Kingdon de l'agenda gouvernemental, « *la liste des sujets ou problèmes auxquels les acteurs gouvernementaux et les personnes à proximité du gouvernement accordent une sérieuse attention à un moment donné* », dans John Kingdon, *Agendas, Alternatives and Public Policies*, New-York, Harper Collins, 1984.

. Voir aussi les travaux de Philippe Garraud, « Politiques nationales : l'élaboration de l'agenda », *L'Année sociologique*, 1990, ou encore la contribution de Jacques Maillard et Daniel Kübler, « Chapitre 1. Les processus de mise à l'agenda », *Analyser les politiques publiques*. sous la direction de Jacques Maillard et Daniel Kübler, Presses universitaires de Grenoble, 2016.

les parlementaires. Il est un fondement argumentatif ensuite car il fournit un panel riche d'indicateurs statistiques problématisés prêts à l'usage, et non de données simplement descriptives, qui permettent de fonder des problèmes éducatifs identifiés par les parlementaires.

Cette double fonction permet d'interroger le gouvernement dans une démarche prospective. A la différence des évaluations de la DEPP qui sont la plupart du temps sollicitées pour ce qu'elles montrent des problèmes internes d'organisation des moyens alloués au système éducatif, l'évaluation de l'OCDE suscite plutôt des questions d'ordre général sur l'avenir du système éducatif, la maîtrise des compétences clés par les élèves français, ou encore la place du système éducatif français dans le monde. Lors de la publication par l'OCDE tous les trois ans sur des résultats principaux de PISA, les parlementaires interrogent le ministre sur ses réactions au positionnement de la France dans les classements ou à l'évolution des résultats de la France dans les classements. Entre chaque publication, PISA est employé pour relancer des sujets jugés fondamentaux pour l'avenir du système éducatif français, comme les orientations nécessaires pour faciliter la maîtrise de la lecture par les élèves français, pour réduire les inégalités scolaires, ou encore pour revaloriser le métier d'enseignant. L'évaluation de l'OCDE apparaît comme un outil d'expertise qui permet d'éviter toute confrontation idéologique, et qui sert de fondement indiscutable à l'argument qui va suivre. C'est un argument d'autorité, qui est préféré à toute autre évaluation nationale pour illustrer efficacement les questions des parlementaires. Elle « révèle », « démontre », « notifie », comme cela est dit dans plusieurs questions orales, des faits sur l'état et les problématiques rencontrées par le système éducatif français, et cela même si certaines des conclusions – la reproduction voire le renforcement des inégalités sociales par l'école par exemple, ou l'inefficacité du redoublement pour améliorer les résultats des élèves – sont par ailleurs connues depuis longtemps des décideurs politiques et étudiées depuis des décennies par les chercheurs⁴⁶³.

⁴⁶³ Nous reviendrons sur la mise en agenda des inégalités scolaires par PISA dans le chapitre suivant, en étudiant son usage dans les deux grandes réformes éducatives de 2005 et 2013.

La question orale posée en séance de questions au gouvernement par le sénateur socialiste Jacques Bernard Magner à la ministre de l'éducation nationale en avril 2015⁴⁶⁴ est un bon exemple de l'emploi de PISA comme argument d'autorité plus efficace que tout autre pour justifier la mise à l'agenda d'un problème public. Alors qu'il interroge la ministre Najat Vallaud Belkacem sur les mesures à venir dans la réforme du collège pour venir à bout du déterminisme social influant sur les résultats des élèves français, il commence par énoncer l'apport des évaluations nationales comme internationales pour prouver son constat de départ, l'effet de l'origine sociale sur les résultats des élèves. Lorsqu'il s'agit de donner des détails statistiques pour renforcer sa preuve, les seules données citées proviennent de PISA, alors même que la DEPP dispose d'un ensemble d'évaluations contextualisées pouvant permettre de satisfaire aux principes de l'evidence-based policy.

« Nous le savons tous, les évaluations nationales et internationales montrent que le collège actuel aggrave la difficulté scolaire, en particulier pour ce qui concerne les disciplines fondamentales. Ainsi, les études PISA témoignent que, en France, entre 2002 et 2012, les collégiens ont régressé en français, en mathématiques et en histoire-géographie, contrairement à ceux des autres pays de l'OCDE, lesquels, en moyenne, ont progressé. En France, la proportion d'élèves ne maîtrisant pas les compétences de base en français est de 12 % en cours moyen deuxième année et de 25 % en troisième. En mathématiques, cette proportion est de 9 % en cours moyen deuxième année et de 13 % en troisième. »

Le sénateur argumente au sujet de ce qu'il considère être un problème bien connu de ses collègues, et tient néanmoins à soutenir son propos par un argument unique et irréfutable, des chiffres qu'il présente comme issus de PISA. Si personne ne trouve à redire au moment où la question est posée en séance, ces éléments de preuve ne correspondent toutefois pas tout à fait au déroulement de PISA. Aucun cycle n'a eu lieu en 2002, mais en 2001 et 2003. De même, les compétences en histoire-géographie ne font pas l'objet d'évaluations propres comme peuvent l'être celles en maîtrise de la langue ou en mathématiques.

⁴⁶⁴ Question au gouvernement n°0504G de M. Jacques-Bernard Magner, Puy-de-Dôme, socialiste, publiée dans le JO du Sénat du 17 avril 2015.

Certaines questions des tests peuvent être relatives à l'histoire et à la géographie dans un grand ensemble de thèmes regroupés sous la dénomination de « sciences humaines », mais elles ne donnent pas lieu à des indicateurs propres comme les autres disciplines citées par le parlementaire. Difficile de savoir si d'autres évaluations, peut-être même des évaluations nationales comme CEDRE qui traite de l'histoire-géographie comme une discipline évaluée à part entière, ont été mélangées à PISA, ou s'il s'agit simplement de mésinterprétations de l'évaluation de l'OCDE. La force de l'argumentation réside simplement dans la mention de PISA, plus que dans l'exactitude de ce qui est rapporté, puisqu'aucune correction de fond ne sera apportée par la ministre dans sa réponse.

La capacité de PISA à être argument d'autorité a pour effet collatéral de limiter le besoin de se plonger véritablement dans les résultats pour comprendre les conclusions présentées par l'OCDE, et par dérivation l'apprentissage de la lecture d'indicateurs parfois de construction complexe et aux résultats trop rapidement – et parfois mal -interprétés. Alors que les parlementaires sont, en particulier après le cycle 2009, en construction de leur maîtrise des outils de l'évaluation, la compréhension des méthodes statistiques derrière les données qu'ils utilisent n'est pas primordiale. Cela est notamment permis par la stratégie d'usage largement choisie par les parlementaires. L'intensité de référence à PISA dans les échanges en séance, et dans une certaine mesure aux statistiques produites par la DEPP, aurait pu donner lieu à un processus nouveau de rationalisation de l'action publique au sens où les idéologies politiques les plus solidement ancrées auraient pu se trouver réfutées par l'utilisation de certains indicateurs ⁴⁶⁵. En pratique, ce processus de rationalisation n'est pas si systématique. En particulier avec PISA, les indicateurs sont sciemment utilisés dans une démarche de politisation de la connaissance. Si les données utilisées par les parlementaires servent à asseoir une argumentation ou à susciter de nouveaux questionnements, ces pratiques nouvelles – ou du moins d'une intensité nouvelle

⁴⁶⁵ Christian Maroy et Catherine Mangez, « Rationalisation de l'action publique ou politisation de la connaissance ? », *Revue française de pédagogie*, 164 | 2008, 87-90.

- d'évaluation de l'action du gouvernement s'accompagnent toujours des interprétations partisans et des détournements avantageux qui caractérisent le jeu politique d'une démocratie majoritaire⁴⁶⁶. Ce n'est pas la politique éducative et les problèmes publics qu'elle vise à résoudre qui sont produits par les indicateurs, mais l'interprétation des indicateurs qui découle des débats parlementaires.

Plusieurs exemples d'usage intentionnel et coordonné de la politisation des résultats de PISA apparaissent à la lecture des questions-réponses. Le plus emblématique est sans doute la fronde lancée par les groupes parlementaires à droite des hémicycles, au Sénat et à l'Assemblée Nationale, entre fin décembre 2013 et mi-janvier 2014, pour exhorter le ministre à abandonner une mesure de sa réforme visant à revaloriser le métier d'enseignant en améliorant les conditions de travail dans les Zones d'Education Prioritaires notamment, qui risquerait selon eux de réduire les salaires versés aux enseignants en Classe préparatoire aux grandes écoles (CPGE). L'une des questions posées au ministre Vincent Peillon au Sénat est ainsi formulée, et utilise les mêmes arguments que toutes les autres posées à la même période sur le même sujet :

« Alors que les conclusions du classement PISA (Program for International Student Assessment) ont démontré que les inégalités s'étaient accrues ces dernières années, il semble important que des mesures soient prises pour lutter contre cette tendance qui hypothèque l'avenir des jeunes mais aussi celui du pays. Aussi, dans ce contexte, il paraît d'autant plus nécessaire de veiller à ce que les formations reconnues pour leur excellence, comme les classes préparatoires puissent conserver leur attractivité auprès des élèves. »⁴⁶⁷

Le sénateur s'appuie sur de véritables conclusions formulées dans PISA, l'aggravation des inégalités scolaires en France montrée à travers leurs indicateurs, pour en tirer une implication allant à l'opposé de l'analyse faite dans PISA, la pérennisation souhaitable de la concentration d'importants moyens dans les filières jugées d'excellence sans que l'allocation des financements ne soit

⁴⁶⁶ Christian Maroy et Catherine Mangez, *art. cit.*

⁴⁶⁷ Question écrite n°09799 de M. Robert Tropeano, Hérault, RDSE, publiée au JO du Sénat le 19 décembre 2013.

davantage équilibrée entre les types de formation. Si l'indicateur sur l'aggravation des inégalités scolaires en France existe bien dans PISA, il est accompagné d'un autre indicateur regardant la performance générale des élèves d'un pays selon le niveau d'inégalités scolaires. L'évaluation de l'OCDE conclut que les pays dans lesquels il existe deux groupes majoritaires d'élèves, très performants aux tests pour l'un et faiblement performants pour l'autre, et où généralement ces répartitions sont corrélées à des niveaux d'inégalités scolaires forts – c'est le cas de la France - sont globalement moins performants que ceux dont un grand nombre d'élèves obtient des résultats moins dispersés. Eric Charbonnier le résumait ainsi en entretien :

« Il faut regarder la qualité d'un système d'éducation. Si tu es comme en France, moyen avec beaucoup de bons élèves et beaucoup d'élèves en échec scolaire, ça ne va pas. Il vaut mieux être moyen avec une dispersion faible de performance. Il y a la question qui est directement liée de l'équité, est ce qu'on a les mêmes chances de réussite quand on vient d'un milieu défavorisé ou non, quand on est issus de l'immigration ou non. En France la réponse c'est non. »⁴⁶⁸

La conclusion tirée des conclusions de l'OCDE au constat posé par le sénateur Tropeano et ses collègues voudrait donc, à moyens constants, que l'on cherche à mieux former les élèves les plus en difficulté, même si cela implique d'allouer moins de moyens aux filières dites « d'excellence », et même si cela a pour conséquence de faire baisser le niveau des meilleurs. On voit ici un exemple flagrant de la politisation de l'action publique par PISA. L'interprétation partielle des indicateurs PISA est préférée à une analyse globale des conclusions tirées par l'organisation internationale. La profusion d'indicateurs de la publication de l'OCDE permet à la fois de pouvoir choisir quelques données traitant plus ou moins le sujet choisi, et dans le même temps de noyer dans la masse d'information à lire les auditeurs qui pourraient souhaiter vérifier les dires des intervenants. Le nom de l'évaluateur est par ailleurs suffisamment légitime pour servir d'argument unique à la démonstration et ne pas être discuté. Il faut en effet

⁴⁶⁸ Entretien avec Eric Charbonnier, Analyse à la Direction de l'éducation, OCDE, 16 décembre 2016.

noter qu'elle ne trouve pas d'opposition de fond des groupes parlementaires de gauche ou du gouvernement, qui aurait pu être construite à partir d'une lecture moins synthétique des publications des indicateurs de l'OCDE. La réponse du ministre aux diverses interrogations sur ce sujet ne viendra en aucun cas expliciter les conclusions complètes de PISA, mais elle se voudra plutôt rassurante, confortant la stabilité des dispositions qui encadrent le salaire des enseignants en CPGE.

Il est important de souligner ici que les mentions de PISA pour asseoir une argumentation sont davantage le fait de parlementaires à droite de l'échiquier politique qu'à gauche. Néanmoins, on note à la lecture des échanges qu'aucun ministre de quelque gouvernement que ce soit, qu'aucun parlementaire de quel groupe politique que ce soit, ne remet en question la validité des résultats PISA, qu'ils soient correctement présentés ou non. Les réponses des ministres aux attaques parlementaires sur leur bilan, sur l'échec de réformes passées ou l'absurdité de leurs réformes fondées sur un argumentaire issu des résultats de PISA ne portent jamais sur la représentativité des données ou sur la prédominance supposée de la méthode-anglo-saxonne. Elles partent toutes du principe que les résultats des élèves français et de leur évolution dans le temps le cas échéant sont robustes, et se placent dans un registre classique de renvoi de la responsabilité sur les gouvernements passés et les réformes antérieures menées par leur opposition.

« À cet égard, je formulerai une simple remarque : les élèves qui ont été évalués en 2009 au titre de l'enquête PISA étaient alors âgés de 15 ans, ce qui signifie qu'ils sont entrés dans le système éducatif en 1997. Il s'agit donc des enfants des réformes qui ont été menées à la fin des années quatre-vingt et au début des années quatre-vingt-dix ! »⁴⁶⁹

« Elle lui demande quelles conclusions il tire de ces résultats (PISA 2012) et quelles mesures en complément de celles-ci déjà prises depuis juin 2012 il entend initier pour prévenir les effets négatifs résultant des politiques menées depuis dix ans. »⁴⁷⁰

⁴⁶⁹ Réponse du ministre de l'Education nationale Luc Chatel à la question crible thématique (question orale) n°0155C posée par la sénatrice Brigitte Gonthier Maurin, Hauts-de-Seine, Parti Communiste Français, publiée au JO du Sénat le 9 février 2011.

⁴⁷⁰ Question écrite n°48629 posée par la députée Marietta Karamanli, Sarthe, Socialiste, publiée au JO de l'Assemblée Nationale le 4 février 2014.

PISA se voit conférer unanimement une position d'autorité qui donne à ses publications une force performative : l'énonciation des problèmes identifiés les place dans le registre de la réalité⁴⁷¹. En résumé, « si PISA le dit, ou plutôt s'il l'écrit, c'est que c'est vrai », ou du moins l'évaluation est suffisamment légitime pour ne pas engendrer de remise en cause des conclusions qui découlent des indicateurs produits et avoir un intérêt majeur dans le jeu des oppositions politiques. L'itération dans le temps des références à PISA⁴⁷², souvent sur des thématiques récurrentes prises isolément de l'ensemble des rapports (le classement de la France, la baisse du niveau des élèves, le renforcement des inégalités scolaires, l'inefficacité du redoublement, ...), c'est-à-dire leur répétition dans des circonstances similaires, renforce la production de l'effet souhaité⁴⁷³. Cet effet performatif est rendu possible par la position d'évaluateur reconnue constitutionnellement des parlementaires, et donc légitime aux yeux de l'auditoire ou des lecteurs des questions posées au gouvernement. Le pouvoir de mise à l'agenda de PISA est donc doublement dépendant des conditions contextuelles⁴⁷⁴ dans lequel il est exercé. D'abord, PISA est légitime aux yeux des décideurs politiques pour constituer une fenêtre d'opportunité⁴⁷⁵ permettant la mise à l'agenda de ses conclusions. Ensuite, les parlementaires sont légitimes aux yeux de leurs collègues et du ministre pour mettre à l'agenda de l'action publique en éducation les conclusions de PISA. Cette conclusion sur la toute-légitimité

⁴⁷¹ Voir l'ouvrage pionnier de J.L. Austin, *How to Do Things with Words*, edited by J.O. Urmson and M. Sbisà, Oxford : Oxford University Press, 2nd édition : 1976 ; traduction française de G. Lane, *Quand dire c'est faire*, Paris : Editions du Seuil, 1970 ; réédité avec une postface de F. Récanati dans la collection « Points-essais », 1991.

⁴⁷² Voir les travaux de Jacques Derrida sur l'itérabilité du signe, et notamment Jacques Derria, *Limited Inc.*, Galilée, Paris, 1990.

⁴⁷³ Franck Aggeri, « Qu'est ce que la performativité peut apporter aux recherches en management et sur les organisations. Mise en perspective théorique et cadre d'analyse », *M@n@gement*, 1, vol.20, 2017.

⁴⁷⁴ Bruno Ambroise, *Les conditions de la performativité et l'efficacité des actes de parole – réponse à J. Butler*, Mars 2007, Poitiers.

⁴⁷⁵ John Kingdon, *Agendas, Alternatives and Public Policies*, 1984. Pour une synthèse des travaux sur le concept, voir l'article de Pauline Ravinet, « Fenêtre d'opportunité », Laurie Boussaguet éd., *Dictionnaire des politiques publiques. 3e édition actualisée et augmentée*, Presses de Sciences Po, 2010.

argumentative de PISA, parfois en méconnaissance – ou du moins en mésusage - de ce que l'évaluation dit du système éducatif français, laisse bien entrevoir par ailleurs les difficultés rencontrées par la DEPP pour mener à bien sa stratégie d'améliorer la compréhension par les décideurs politiques français d'une évaluation internationale à la méthode complexe et aux productions foisonnantes. La sortie des publications PISA combinée la parole parlementaire émise dans un contexte de contrôle de bonne orientation de l'action publique constitue une opportunité de placer l'évaluation de l'OCDE au centre du débat. Par-là, elle rend visible les problèmes publics, mis en avant par une organisation internationale de plus en plus relayée dans les espaces médiatiques et politiques⁴⁷⁶, et dans un contexte, nous le verrons ensuite, d'internationalisation des enjeux des politiques éducatives nationales.

2) Comparer: usage différencié des évaluations et internationalisation des défis de l'éducation nationale

a) L'utilisation de la comparaison dans le contrôle continu de l'action de l'Etat par les parlementaires

Comme nous l'avons vu au chapitre précédent, la DEPP a tenté dès les premiers cycles PISA de réfréner les usages des résultats de l'évaluation internationale sous la forme de palmarès ou de classements dans lesquels les résultats des élèves français seraient comparés à ceux des élèves des autres pays participants. Sans nier l'intérêt de la démarche de comparaison, la direction du ministère a tenté de replacer la focale de la comparaison spatiale, entre pays, vers la comparaison temporelle, du pays par rapport à lui-même. Nous avons regardé au sein des questions citant PISA la répartition des usages entre deux critères : l'utilisation de PISA pour le classement qu'il donne des résultats des élèves

⁴⁷⁶ Sur l'effet de sélection des problèmes publics dans les arènes publiques, voir notamment Charles Bosk et Stephen Hilgartner, « The Rise and Fall of Social Problems : A Public Arena Model », *American Journal of Sociology*, vol.94, n°1, 1988.

français ou de la France, et son utilisation pour justifier des préoccupations de politique nationale, sans mention de classement. Un critère intermédiaire a également été observé, celui d'une double utilisation des résultats français à PISA, visant à la fois à interroger sur une problématique du fonctionnement du système français souvent dans une perspective temporelle (le problème existe de longue date et les résultats confirment son existence, ou le problème s'est aggravé et les résultats le montrent, ...), et à mettre en avant le classement de la France dans l'évaluation de l'OCDE.

Cette étude nous permet deux choses. D'abord, objectiver l'importance de la fonction de comparateur international de l'outil PISA par rapport à une fonction explicative qui serait resserrée sur ce que l'évaluation dit du système éducatif dans son cadre national. Ensuite, d'établir dans quelle mesure la stratégie de la DEPP visant à réduire l'intérêt explicatif de l'utilisation de palmarès internationaux a pu réussir, en parallèle de sa tentative d'apporter un nouvel outil adapté à la conception française des résultats attendus de l'éducation, CEDRE. Dans les faits, si l'on regarde la façon dont PISA est mobilisé dans les questions des parlementaires depuis ses premières publications, l'usage de la comparaison des performances attribuées au système français reste toujours formulé sous la forme d'une comparaison de la France avec les autres pays participants, même si la mention pure du palmarès PISA est de moins en moins sollicitée dans les échanges de questions au gouvernement au fil des ans.

A l'Assemblée nationale et au Sénat, près de la moitié des questions posées par les parlementaires abordent PISA en présentant une comparaison internationale des résultats, sans que cela ne soit accompagné de la formulation d'un problème identifié ou suspecté de fonctionnement du système éducatif français. Il s'agit par exemple d'une demande de réaction du ministre suite à la publication des résultats PISA et au classement jugé mauvais de la France, ou de l'expression de l'inquiétude que provoquent les résultats des élèves français par rapport aux autres.

Un deuxième type d'usage est principalement employé après 2010 compte-tenu de la disponibilité de plusieurs séries de données antérieures comparables.

Plus de 10% des questions utilisent non seulement le classement de la France par rapport aux autres pays participants, mais s'accompagnent aussi d'une comparaison dans le temps du rang obtenu par la France, ainsi que de pistes de travail. Le classement de la France est mauvais, mais en plus il empire, tandis que les autres s'améliorent tandis que la France stagne, et il faut faire quelque chose pour arrêter la tendance à la baisse. Certains, interrogent le ministre sur l'évolution des méthodes d'apprentissage de la lecture et sur les mauvais résultats de la méthode syllabique employée en France, parlent ainsi de PISA pour appuyer leurs propos en insistant sur le fait que la France recule dans tous les classements de l'évaluation. D'autres encore alertent le ministre sur la baisse de performance scolaire des élèves, et citent PISA et les résultats des élèves au fur et à mesure des cycles pour preuve de cette baisse généralisée du niveau qu'ils dénoncent.

Le troisième usage de PISA, qui apparaît également majoritairement après décembre 2010, est celui d'un argument permettant de justifier un besoin de réforme sans mention de classement, uniquement fondé sur des problématiques de politique nationale qui s'aggravent de cycle en cycle. Les résultats PISA témoignent par exemple d'un danger pour le modèle social français en ce qu'ils montrent une baisse du niveau scolaire qui s'accompagne de l'augmentation continue du nombre de décrocheurs. Ils incitent à travailler davantage sur la prévention de l'échec scolaire, sur la lutte contre l'illettrisme, ou encore de recalibrer urgemment les mesures de lutte contre les inégalités scolaires.

Même si l'accroissement de ce troisième usage des données PISA est concomitant d'un accroissement de l'usage des travaux de la DEPP, l'attrait de la comparaison internationale semble difficile à contrecarrer pour la direction du ministère. Les comparaisons internationales qui fondent largement les arguments portés par les parlementaires portent déjà en elles, par la façon dont sont construits les indicateurs PISA, une dimension explicative des enjeux à venir. Dans la conception de PISA portée à la fois par Andreas Schleicher et par les responsables de la communication autour des résultats PISA, le classement n'est

en effet qu'une étape d'un processus d'évaluation que la France a du mal à mettre en œuvre jusqu'au bout.

« Avoir un classement, c'est vrai que c'était la base de la réflexion. Avoir des pays qu'on ne s'attendait pas à trouver en tête des classements, comme la Finlande par exemple, quand on a vu Finlande premier, on s'est tous regardés et on a eu des réflexions pour comprendre pourquoi. On a vraiment eu un chemin de réflexion qui était de passer de la notion de performance à la notion de qualité. Qu'est ce qui a fait que le système était de qualité. Une des dimensions c'est l'équité sociale. On ne peut pas avoir autant d'inégalités sociales que celles que tu as en France par exemple. »⁴⁷⁷

Selon Eric Charbonnier, analyste à la Direction de l'éducation de l'OCDE, la performance que l'on peut lire à travers le classement des résultats PISA mène vers la quête de qualité, et n'est pas un objectif en soi. Elle implique un travail d'analyse pour déterminer des pistes d'action de politique éducative. La DEPP, chargée de rendre compte des résultats des élèves français à PISA, ne détient pas dans ses missions officielles le rôle d'analyse et de formulation de recommandations d'action. Celui-ci reviendrait théoriquement davantage à la Dgesc, mais en pratique les notes analytiques sont souvent rédigées par les directeurs et directrices successifs de la DEPP directement aux nouveaux ministres à leur arrivée en poste. Les équipes de la DEPP ne sont par ailleurs pas identifiées par les décideurs politiques et notamment les parlementaires comme étant partie prenante de PISA et donc potentiellement des interlocuteurs clés pour en comprendre le fonctionnement et les implications. Personne au sein du ministère n'a finalement le rôle officiel de se saisir systématiquement des résultats PISA pour les traduire en propositions d'action publique, alors même que leur utilisation pour le contrôle de l'action de l'Etat est de plus en plus fréquente. La DEPP, la mieux à même de pouvoir fournir une analyse problématisée de la masse de données fournies par l'OCDE compte tenu de sa maîtrise technique de l'évaluation et de son immersion constante dans les réunions de travail organisées par l'organisation internationale, n'est pas sollicitée pour compléter son travail

⁴⁷⁷ Entretien avec Eric Charbonnier, Analyste à la Direction de l'éducation, OCDE, 16 décembre 2016.

d'opérateur de PISA en France. PISA devient un outil de pilotage de l'action publique en éducation sans véritable pilote, ce qui laisse toute latitude à l'OCDE de déterminer à chaque cycle les implications qu'elle juge prioritaire au vu des résultats de chaque pays participant. Des « country notes » d'une quinzaine de pages sont ainsi rédigées à chaque cycle, pour mettre en avant les résultats saillants de plusieurs pays. Si ces rapports ne contiennent pas à proprement parler de recommandations de réformes, ils contribuent à normaliser certains objectifs des politiques éducatives : la lutte contre les inégalités scolaires, les inégalités de genre ou les inégalités liées à l'origine étrangère, les stéréotypes de genres portant sur les métiers, le soutien ressenti par les élèves de la part de leurs enseignants, l'absence de « mentalité de croissance » c'est-à-dire la croyance des élèves dans le fait que l'intelligence peut se développer au cours de la vie, ... Les *country notes* sont accompagnées par d'autres publications cette fois bien plus explicites quant aux recommandations que l'OCDE formule à destination des décideurs politiques français. C'est le cas notamment de la publication *Perspective des politiques de l'éducation* mise en place à échéance régulière à partir de 2013, accompagnée elle aussi de plusieurs rapports spécifiques par pays⁴⁷⁸, dans le cas de la France en 2014 et 2020.

L'OCDE est alors un diffuseur de priorités de politiques publiques, sans que l'organisation ne produise réellement et clairement un modèle unique à suivre issu des résultats de PISA. Chaque pays est libre de tirer ses propres conclusions sur le modèle à suivre pour améliorer le fonctionnement de son système éducatif, si tant est qu'il souhaite améliorer ses résultats à PISA. Les recommandations ne sont évidemment pas coercitives, puisque l'OCDE ne dispose pas de ce pouvoir d'imposition de réformes à ses pays membres. L'organisation internationale se situe davantage ici dans le cadre de la diffusion de politiques publiques telle que définie par Busch et Jörgens dans leur étude des types de transfert de politiques publiques, qui considèrent la diffusion comme le « processus par lequel les innovations en politiques publiques sont

⁴⁷⁸ Voir par exemple le dernier volume, OCDE, *Perspectives des politiques de l'éducation – France*, juin 2020.

communiquées au système international et adoptées volontairement par un nombre croissant de pays au fur et à mesure du temps ». ⁴⁷⁹ La diffusion est envisagée à une échelle internationale et correspond à l'intégration progressive dans les discours de pilotes de l'action publique de l'usage de concepts, d'outils, de méthodes ou de la définition d'objectifs à atteindre. Par la multiplicité des problématiques qu'elle traite, l'OCDE porte avec PISA un panier de sujets dont les gouvernements peuvent se saisir au gré de leurs priorités politiques, et dont les parlementaires peuvent se saisir au gré des préoccupations qu'ils souhaitent porter au-devant du débat en hémicycle. Elle porte également une conception propre de la qualité de l'éducation, qui est reprise ou non par les pays. L'usage de la comparaison des données et des conclusions de PISA lui confère une « sur-référentialité » de fait, si l'on reprend le concept formulé par Olivier Ducrot ⁴⁸⁰. La répétition tous les trois ans de l'évaluation et la publication constante de rapports thématiques traitant de problématiques de politique éducative spécifique ou de pays spécifiques rend non seulement PISA évocable à échéances régulières, mais elle permet à PISA de faire référence à lui-même en encourageant la comparaison avec les éditions précédentes. Le « produit » PISA gagne en quelque sorte en légitimité en se citant lui-même à chaque nouvelle publication, et en étant cité dans le cadre national de plusieurs dizaines de pays participant à l'évaluation, amenés par ailleurs à échanger dans d'autres forums internationaux et à y faire entrer les enjeux de PISA. Si, par construction de ses missions d'évaluation à un niveau national, la DEPP ne peut bien évidemment pas entrer sur le terrain des comparaisons internationales, on peut faire l'hypothèse que le produit CEDRE, identifiable car conceptualisé en produit et bénéficiant de l'aura de PISA en ce qu'il est régulièrement cité en contrepoint de l'évaluation de l'OCDE, a sans doute depuis quelques années profité progressivement du même mécanisme de

⁴⁷⁹ Per Olof Busch et Helge Jorgens, « The International sources of policy convergence : explaining the spread of environmental policy innovations », *Journal of european public policy*, 12 octobre 2005.

⁴⁸⁰ Olivier Ducrot, « Illocutoire et performatif », *Dire et ne pas dire*, Paris, Hermann 3^e édition : 1991.

sur-référentialité pour gagner en visibilité comme en légitimité auprès des parlementaires.

L'échec de la DEPP de réduire l'intérêt du palmarès proposé par l'OCDE, semble évident, car finalement les classements présentés dans les publications PISA sont utilisés largement par les décideurs politiques, même lorsque la comparaison des résultats des élèves français par rapport à eux-mêmes devient possible. Cet attrait irrésistible pour la comparaison internationale s'inscrit dans un contexte structurant d'internationalisation des enjeux de l'éducation nationale à laquelle la France non seulement adhère pleinement mais dans lequel elle joue un rôle moteur.

b) Le renforcement de l'internationalisation des enjeux de l'éducation : un contexte défavorable à la mise à l'agenda des travaux de la DEPP.

La majorité des parlementaires comme des ministres commentent en séance les résultats obtenus par « la France » par rapport aux autres pays participants. Cette substitution, dans le discours, des résultats des *élèves français* par ceux de la *France* confèrent à PISA un rôle d'évaluation des systèmes éducatifs et donc des politiques publiques qui les créent et les modifient – et non d'évaluation des acquis des élèves de 15 ans comme son nom l'indique pourtant. Le déplacement de l'objet de l'évaluation s'inscrit dans un cadre d'eupéanisation des enjeux de l'éducation. Il vient donc acter la reconnaissance par les parlementaires et le gouvernement du bienfondé de l'évaluation du système éducatif français par des évaluateurs internationaux.

La participation de la France à des démarches internationales de réflexion sur les politiques éducatives nationales ne date évidemment pas des années 2000⁴⁸¹. Néanmoins, la stratégie européenne de Lisbonne est lancée en 2000 et renforce considérablement l'intérêt des décideurs politiques français pour les outils leur permettant d'agir vers l'atteinte de ses objectifs. Les volets d'action de

⁴⁸¹ Voir chapitre 2.

la stratégie européenne sont dirigés explicitement vers la préparation de la démarche qui devra prévaloir pour favoriser le développement économique de l'Europe du futur, grâce d'abord à un travail coordonné sur « la transition vers une économie compétitive, dynamique et fondée sur la connaissance »⁴⁸² (volet 1). L'aboutissement de ce premier volet est associé à celui du second, la modernisation d'un « modèle social européen » (volet 2). Plusieurs concepts de la stratégie de Lisbonne ramènent aux principes qui dirigent l'élaboration des cycles PISA au fur et à mesure des années. Le travail engagé par les pays de l'Union européenne vers la construction d'une entité politique et économique forte est étayé par l'application de méthodes issues du management, et notamment la Méthode ouverte de coordination (MOC), mettant en œuvre un *benchmark* (ou parangonnage en Français) des performances des pays membres, à partir de la définition d'un cadre commun d'objectifs opérationnalisés en indicateurs uniques pour tous⁴⁸³. PISA de l'OCDE s'inscrit dans cette démarche pour au moins deux raisons majeures. Il permet bien sûr d'une part aux pays européens d'avoir clé en main un cadre d'analyse commun pour mesurer les performances de leurs systèmes éducatifs, avec en plus un point de comparaison avec des partenaires économiques n'appartenant pas à l'Union européenne mais complètement intégrés dans la course à la compétitivité internationale, et notamment la compétitivité des travailleurs formés par les différents systèmes éducatifs. D'autre part, en s'inscrivant dans une perspective prospective et mouvante des attendus de l'éducation qu'il mesure à chaque cycle, afin de mieux correspondre à ce que l'OCDE estime nécessaire de pouvoir faire une fois les futurs actifs entrés sur un marché du travail en perpétuel changement, PISA offre aux pays européens un outil qui s'intègre dans leur stratégie commune d'un dynamisme économique fondé sur la connaissance.

En parallèle, la DEPP, qui certes travaille à PISA mais dont les analyses portent sur le cadre français avec peu de comparaisons internationales, ne fournit

⁴⁸² Conclusions du Conseil Européen de Lisbonne, 23 et 24 mars 2000.

⁴⁸³ Isabelle Bruno, « La déroute du « benchmarking social ». La coordination des luttes nationales contre l'exclusion et la pauvreté en Europe », *Revue Française de Socio-Economie*, La Découverte, n°5, 1, 2010.

pas cet outil au cœur de la nouvelle stratégie concertée des pays européens. Il est difficile alors pour les usagers des évaluations en éducation, au premier rang desquels les fabricants de la loi, membres du gouvernement ou parlementaires, de se référer à la DEPP pour justifier leurs initiatives en matière de réformes ou la visibilité de celles-ci auprès des partenaires européens de la France. La popularité de PISA et les conditions propices à sa réception par les parlementaires français s'expliquent quant à elles en partie par la construction d'un référentiel de discours particulier et commun, celui de l'euphorisation des enjeux éducatifs, qui engendre une harmonisation des champs des possibles en politique éducative dans les pays qui y participent. Le référentiel dans lequel évolue PISA devient à la fois compréhensible et plausible⁴⁸⁴ pour les acteurs politiques des pays européens, et donc pour la France.

Dans ce cadre de l'internationalisation des problématiques éducatives au niveau européen, l'évaluation PISA devient un enjeu de visibilité au niveau international, ainsi qu'un outil permettant d'évaluer les ressources en « capital humain ». PISA s'inscrit totalement dans cet objet récent de la théorie économique ayant l'oreille des gouvernements depuis le début des années 2000⁴⁸⁵, promouvant au rythme de ses publications une logique de comparaison à venir des ressources humaines disponibles et efficaces dans chaque pays, et donc de son « gisement de productivité »⁴⁸⁶. Par la communication élaborée qu'elle met en place progressivement⁴⁸⁷, l'OCDE installe avec PISA une stratégie de pression par les pairs⁴⁸⁸ dans le champ non seulement de la politique éducative portant sur les années de scolarité obligatoire, mais aussi dans celui de la politique de formation et de l'emploi. Cette pression par les pairs est tout à fait consciente chez les parlementaires qui la considèrent positivement comme un effet

⁴⁸⁴ James E. Anderson, *Public Policymaking*, Praeger, 1975.

⁴⁸⁵ Elisabeth Chatel, « L'évaluation de l'éducation dans le contexte de la mondialisation : le cas français », *Tréma*, 40, 2013.

⁴⁸⁶ Marc Gurgand, *Economie de l'éducation*, La Découverte, Paris, 2005.

⁴⁸⁷ Nathalie Mons, « Policy Responses to PISA : a comparative investigation across OECD countries », Keynote speech for the 14th International Conference of the Association for Educational Assessment, *International Surveys, Policy Borrowing and National Assessment*, Université de la Sorbonne, Paris, 2013.

⁴⁸⁸ Xavier Pons, *Evaluer l'action éducative. Des professionnels en concurrence*, PUF, 2010, p.96.

« *d'apprentissage et d'émulation* ». La formule est notamment reprise d'un article de Jean Pisani-Ferry et André Sapir⁴⁸⁹ et repris dans le cadre des travaux menés par le Comité d'évaluation et de contrôle à l'Assemblée nationale⁴⁹⁰ sur le lien entre stratégie de Lisbonne et état de l'économie française.

« Des effets d'apprentissage et d'émulation

La Stratégie de Lisbonne emporterait également un effet d'apprentissage mutuel. Dans cette perspective, les États membres pourraient tirer des conclusions de l'évaluation du succès des réformes réalisées dans les autres États membres et sélectionner ainsi les meilleures pratiques.

Dans leurs travaux précités, MM. Jean Pisani-Ferry et André Sapir mettent en évidence que mener des politiques économiques de manière simultanée pourrait, sous certaines conditions, amener les États à apprendre des succès, des difficultés, voire des échecs des autres États membres. Cet effet d'apprentissage, qui serait particulièrement marqué s'agissant des politiques visant à réformer les marchés du travail, serait cependant conditionné à la transparence des processus, qu'il s'agisse des objectifs, des performances et des méthodes d'évaluation. Il ne fonctionnerait en effet que si le système garantit une information objective et détaillée sur les efforts fournis, les performances réalisées et les gains retirés par chaque État.

Cet effet pourrait aussi s'accompagner d'une autre conséquence positive, celle d'une mobilisation par l'émulation : chaque État obtiendrait une connaissance précise de sa situation relative par rapport à celle de ses voisins, et tenterait alors de les surpasser dans la mise en œuvre formelle et effective des politiques considérées. MM. Pisani-Ferry et Sapir indiquaient que cet effet d'émulation relative avait pu être constaté s'agissant des études PISA sur les performances des systèmes scolaires, études comparatives menées dans le cadre de l'OCDE. »

⁴⁸⁹ Jean Pisani-Ferry, André Sapir, « Last Exist to Lisbon », *Bruegel Policy Contribution*, mars 2006.

⁴⁹⁰ Rapport d'information n°4364 de MM. Philippe Cochet et Marc Dolez sur l'évaluation des incidences de la Stratégie de Lisbonne sur l'économie française, comité d'évaluation et de contrôle des politiques publiques, 16 février 2012.

On comprend en lisant le rapport des députés Philippe Cochet⁴⁹¹ et Marc Dolez⁴⁹² l'effet de pression qui a pu s'exercer depuis le lancement de PISA, une fois passé le premier temps d'évitement et de critiques méthodologiques, en voyant certains proches voisins européens engager un travail concret d'amélioration de leur système éducatif en conséquence des premiers résultats de l'évaluation⁴⁹³. Andreas Schleicher affirmait clairement en entretien ces objectifs d'exemplarité et de pression par les pairs alloués à PISA, et dont la réussite était permise par la possibilité, avec un peu de patience, de faire des comparaisons temporelles entre les résultats de différents cycles.

« Et probablement ce qui s'est passé pour les autres pays c'est qu'au premier cycle de PISA ils ont eu juste un classement de pays. La deuxième et la troisième fois ils ont pu vraiment voir qui avait progressé et qui n'avait pas progressé. Et je pense que le fait qu'il n'y ait pas eu de grand changement en France, alors que des pays comparables ont fait des progrès significatifs, ça a dû créer un questionnement supplémentaire : « pourquoi ces pays fonctionnent bien ? Pourquoi ont-ils fait des progrès plus rapidement que nous ? » Je suis certain que cela est très lié à la dimension temporelle de PISA». ⁴⁹⁴

La logique de classement, dont Schleicher se défend d'avoir voulu faire un élément phare et incontournable de PISA, est en réalité centrale pour que la pression par les pairs fasse effet. L'objectif de PISA étant de donner à voir non seulement les bonnes pratiques dans un contexte donné mais aussi d'inciter les pays à davantage échanger et partager leurs essais de politique éducative⁴⁹⁵, le classement et la popularisation à l'échelle mondiale de ce classement par ce biais,

⁴⁹¹ Député Les Républicains du Rhône de 2002 à 2017.

⁴⁹² Député du Nord de 1988 à 1993 puis de 1997 à 2017. Ancien membre du Parti Socialiste, il a contribué à la fondation du Parti de Gauche avec Jean-Luc Mélançon.

⁴⁹³ Bien évidemment l'Allemagne dans le cadre du choc PISA déjà mentionné, mais aussi par exemple l'Espagne, le Portugal, ou la Pologne, qui ont engagé des réformes structurelles de leurs systèmes éducatif et d'enseignement supérieur à la suite des premières évaluations PISA.

⁴⁹⁴ Entretien avec Andreas Schleicher, Directeur de la Direction de l'éducation, OCDE, 16 janvier 2017. Traduction par nous : « *And you know surely what other countries have done is the first time you just had a ranking of countries. The second and third time you actually see who is progressing and who's not. And I think that the fact that there was not a great change in France, when similar countries have made significant progress, I think that also create an additional "why are these countries doing well, why are they making faster progress?" So I think the time dimension has to do a lot with that.* »

⁴⁹⁵ Voir le chapitre 2 et le mythe créateur de PISA raconté par Andreas Schleicher.

efficace du point-de-vue de la communication, est le moyen le plus efficace de l'atteindre.

A la lecture des interventions des parlementaires français, l'effet d'émulation semble avoir été plus immédiat que l'effet d'apprentissage. Les interventions portant sur l'image de la France auprès de ses partenaires européens et internationaux, mettant en parallèle les classements PISA obtenus et le positionnement de la France en termes de puissance économique, sont régulières tout au long de la période que nous étudions. La formulation de la question du sénateur Alain Marc à la Ministre Najat Vallaud-Belkacem en est un bon exemple.

« Nous constatons aujourd'hui une baisse de la performance scolaire de nos élèves. Le classement international du programme PISA nous place au vingt-sixième rang ; personne ne peut nier que ce classement est insatisfaisant pour la sixième nation du monde.»⁴⁹⁶

Cette intervention du sénateur Alain Marc n'est pas sans rappeler l'anecdote racontée en entretien par Bruno Trosseille, ancien chef du Bureau de l'évaluation des élèves, concernant l'appel reçu par l'équipe PISA-DEPP provenant du Ministère des Affaires Etrangères et exigeant que « *les résultats français s'améliorent* »⁴⁹⁷. Par cette formulation impérative, les équipes travaillant sur PISA à la DEPP se trouvent engagées dans un travail d'évaluation à la portée plus large que le champ de la politique éducative et de la politique nationale. PISA se positionne comme un outil aux implications de politiques étrangères fortes, dont les résultats doivent certes servir à accompagner l'amélioration du fonctionnement du système éducatif, mais qui a aussi à court terme des implications sur l'image internationale des pays participants, aussi bien du point-de-vue économique que de l'attractivité de leur système d'enseignement et de formation. Des inquiétudes avaient ainsi été exprimées par les parlementaires quant aux effets possiblement dévastateurs de PISA sur l'attractivité des

⁴⁹⁶ Question d'actualité au gouvernement n°1109G de M. Alain Marc, Aveyron, Les Républicains, publié au JO du Sénat le 18 janvier 2017.

⁴⁹⁷ Entretien avec Bruno Trosseille, Directeur du bureau des évaluations des élèves, DEPP, 5 janvier 2016.

formations d'enseignement supérieures françaises, aussi bien sur les jeunes français tentés d'aller plutôt à l'étranger, que sur les jeunes étrangers tentés d'aller suivre les formations dispensées dans des pays mieux classés dans PISA. La question avait été posée directement à Pascal Faure, directeur général de la Direction générale de la compétitivité, de l'industrie et des services (DGCIS) au ministère de l'Économie, du redressement productif et du numérique, lors de son audition dans le cadre du rapport d'information sur « *l'exil des forces vives de France* »⁴⁹⁸. Il s'était voulu rassurant, sans pour autant aller plus loin que l'expression d'une impression générale :

« Très souvent, sur ce sujet, on manque d'outils objectifs pour appréhender la réalité, ce qui appelle, encore une fois, une certaine prudence. En effet, quand on creuse, on s'aperçoit que la réalité ne correspond pas toujours à ce qu'on lit par ailleurs. Je n'ai pas de chiffres sur le nombre d'Européens entrés sur notre territoire, mais je peux essayer de vous en trouver.

En revanche, on arrive à mesurer la durée de séjour en France des étrangers qui rentrent pour des raisons professionnelles pour la première fois : au bout d'un an, environ 40 % sont repartis, au bout de trois ans, 50 % d'entre eux, et au bout de sept ans, plus de 60 %. Le temps de séjour médian est donc à peu près de trois ans. On doit avoir des chiffres semblables pour les Français allant dans des pays comparables à la France.

La France dispose d'avantages majeurs, liés non seulement à la taille de son marché et à la qualité de ses infrastructures, mais aussi à celle de son système de soins et de son système éducatif – facteur important pour les cadres, qui décident du lieu d'implantation et viennent généralement avec leur famille. Or ces deux systèmes sont bien perçus à l'étranger, nonobstant les études PISA. »

Là encore les statistiques répondent aux statistiques, pour fournir des preuves bien qu'un peu éloignées sur l'attractivité persistante de la poursuite d'une carrière professionnelle en France malgré une probable mauvaise presse du fait

⁴⁹⁸ Rapport fait au nom de la commission d'enquête sur l'exil des forces vives de France, présidé par Luc Chatel, Rapporteur Yann Galut, 8 octobre 2014.

des résultats à PISA. Le lien fait régulièrement entre l'évaluation de l'OCDE, qui porte sur les élèves de 15 ans, et l'attractivité du marché du travail national renforce l'hypothèse selon laquelle la puissance argumentative de PISA, contrairement aux travaux de la DEPP, bénéficie du contexte d'internationalisation des enjeux d'enseignement supérieur et de l'emploi qui se renforce significativement dès le début des années 2000 avec la formalisation de la stratégie de Lisbonne. La dynamique européenne a eu pour effet de standardiser les outils de mesure de l'action publique en éducation⁴⁹⁹, dans le but d'atteindre des objectifs donnés certes dans la sphère éducative, mais avec l'objectif en second jour d'apporter des réponses aux évolutions du marché du travail. La stratégie de Lisbonne s'est ainsi retrouvée doublée par la mise en place de la stratégie européenne pour l'emploi, qui prévoit une série d'objectifs sur l'emploi des jeunes, sur la formation continue et sur le développement du capital humain. C'est bien là pour l'OCDE l'un des messages centraux qu'elle porte avec le véhicule PISA. L'approche des résultats de l'éducation en termes de compétences non-disciplinaires plus qu'en termes d'apprentissage d'un programme, même construit avec une vision prospective des enjeux du marché du travail à venir, est présentée comme la mieux adaptée à l'incertitude inhérente aux formes futures des emplois des pays de l'OCDE. Et en corollaire, une meilleure adaptation aux formes futures des emplois permet d'améliorer la compétitivité des pays, véritable leitmotiv de l'organisation depuis sa création⁵⁰⁰.

Le sujet est central pour les parlementaires qui font du lien entre les résultats PISA et l'état du marché du travail futur constitué des élèves d'aujourd'hui, un vrai thème d'intervention et d'inquiétude récurrent. Les mauvais classements des résultats des élèves de quinze ans sont un double objet

⁴⁹⁹ Romuald Normand, *Gouverner la réussite scolaire. Une arithmétique politique des inégalités*. Berne : Peter Lang et Lyon : ENS de Lyon, 2011.

⁵⁰⁰ OCDE, *Analyse des politiques d'éducation*. Paris : OCDE, 1997 ; OCDE, *Analyse des politiques d'éducation*. Paris : OCDE, 2005 ; OCDE, *L'éducation aujourd'hui. La perspective de l'OCDE*. Paris : OCDE, 2013) ; Fazal Rizvi & Bob Lingard, « The OECD and Global Shifts, in Education Policy », dans Robert Cowen & Andreas M. Kazamias (ed.), *International Handbook of Comparative Education*, Dordrecht : Springer Netherlands, 2009 ; Thibaut Lauwerier, « L'éducation au service du développement. La vision de la Banque mondiale, de l'OCDE et de l'UNESCO. » *L'éducation en débats : analyse comparée*, 2017, vol.9, pp 43-58.

d'inquiétude car cela est susceptible d'entraîner l'entrée en formation supérieure d'étudiants aux compétences insuffisantes ou inférieures à celles des étudiants d'autres pays en concurrence avec la France, et l'entrée sur le marché du travail d'actifs aux compétences moins adaptées aux évolutions des formes d'emploi que d'autres actifs formés dans d'autres pays. Plusieurs rapports parlementaires abordent ces thématiques et donnent lieu pour certains à des débats en séance sur le sujet. Le député Emeric Bréhier⁵⁰¹, rapporteur du rapport d'information déposé par la commission des affaires culturelles et de l'éducation de l'Assemblée nationale en conclusion des travaux de la mission sur les liens entre le lycée et l'enseignement supérieur en juillet 2015, s'inquiète de voir arriver une « cohorte » d'étudiants aux résultats PISA « très dégradés ⁵⁰²». Le rapport, rendu à la même époque, portant sur l'impréparation des élèves français au monde du travail, indique quant à lui en gras la proportion importante de futurs actifs mal préparés à l'entrée sur le marché du travail, directement tirée des données PISA pourtant concernant les élèves de 15 ans.

« Le progrès technologique rapide a transformé les besoins sur le marché du travail, alors même que les entreprises évoluent sur un marché mondial où la concurrence fait rage. Au total, au niveau mondial, ce sont les travailleurs les plus instruits ou possédant des compétences spécifiques qui sont les plus demandés, tandis que les travailleurs peu qualifiés risquent davantage de voir leur emploi automatisé.

Or, selon un rapport récent de l'OCDE, **un élève sur cinq ne serait pas préparé au monde du travail** ⁵⁰³ Mesurant les performances dans l'enquête PISA 2012 de près de 85 000 jeunes de 15 ans dans 44 pays de l'OCDE, l'organisation a jugé que « les élèves âgés de 15 ans aujourd'hui qui ont une capacité médiocre en résolution de problèmes seront demain des adultes qui auront des difficultés à trouver et garder

⁵⁰¹ Emeric Bréhier est député de Seine et Marne de juin 2012 à juin 2017. Docteur en science politique dont la thèse s'intitule *Les revues politiques de la gauche non communiste de 1958 à 1986* sous la direction de Pierre Avril à l'IEP de Grenoble, il s'engage en 2000 au Parti Socialiste. Il en est exclu en 2018 suite à son refus de soutenir Benoît Hamon, candidat officiel à l'élection présidentielle en 2017, et son soutien à La République en Marche. Il est désormais directeur de l'Observatoire de la vie politique de la Fondation Jean Jaurès.

⁵⁰² Rapport d'information n°2951, déposé par la commission des affaires culturelles et de l'éducation en conclusion des travaux de la mission sur les liens entre le lycée et l'enseignement supérieur, présenté par Emeric Bréhier, député, le 8 juillet 2015, p.144.

⁵⁰³ En gras dans le rapport.

un emploi. Les décideurs et les éducateurs devraient refondre les systèmes éducatifs et les programmes d'enseignement de manière à aider les élèves à développer les capacités de résolution des problèmes dont ils auront de plus en plus besoin dans les économies contemporaines ». Les pays asiatiques sont à cet égard en tête du classement, tandis que neuf pays de l'Union européenne, dont la France, réalisent un score au-dessus de la moyenne de l'OCDE aux côtés du Canada, des États-Unis et de l'Australie : Finlande, Royaume-Uni, Estonie, France, Pays-Bas, Italie, République Tchèque, Allemagne et Belgique. »⁵⁰⁴

Les données de PISA observées au niveau national uniquement ne sont pas de nature à rassurer les parlementaires sur les résultats de la politique éducative française. Néanmoins, il faut noter ici un usage de la comparaison spatiale dans le but de rassurer quant à la position de la France parmi quelques voisins européens au-dessus de la moyenne OCDE, et à un niveau similaire à celui des concurrents-anglo-saxons.

Les questions-réponses parlementaires sont également un moment d'expression de ces inquiétudes, exacerbées par les résultats à PISA :

« Le constat est aujourd'hui édifiant. La France ne cesse de reculer au classement PISA. La maîtrise de la langue française est désastreuse, même chez les étudiants de bon niveau. »⁵⁰⁵

« Cette situation a été aggravée par la dernière majorité qui considérait l'éducation comme un fardeau financier et non un investissement pour l'avenir : la dernière étude PISA a confirmé que, depuis dix ans, les écarts de niveau entre élèves se sont creusés et que le poids des inégalités sociales dans la réussite scolaire s'est alourdi. »⁵⁰⁶

PISA témoigne dans ces exemples non seulement de la capacité des élèves à trouver dans le futur un emploi de qualité non automatisable, mais aussi à rester dans l'emploi. Le lien est alors fait entre PISA, la qualité des formations dispensées dans l'enseignement supérieur vue au prisme des compétences des

⁵⁰⁴ Rapport d'information déposé par la commission des affaires européennes le 4 mars 2015 sur l'emploi des jeunes en Europe, présenté par Philip Cordery, Assemblée nationale.

⁵⁰⁵ Question n°2976 de Monsieur Jean-Louis Costes, Lot-et-Garonne, Les Républicains, publiée au JO de l'Assemblée nationale le 10 juin 2015.

⁵⁰⁶ Réponse du ministre de l'Éducation nationale à la question au gouvernement n°1528 posée par Monsieur Hervé Fréron, Meurthe-et-Moselle, groupe socialiste, radical, citoyen et divers gauche, publiée au JO de l'Assemblée nationale le 16/01/2014.

étudiants qui en seront diplômés, la qualité de la situation de l'emploi en France, et potentiellement les chiffres du chômage qui pourraient être particulièrement mauvais si les actifs n'étaient pas les mieux armés pour rester en emploi durablement.

Alors que l'OCDE est capable de répondre par une seule voix à tous ces enjeux prospectifs de politiques publiques, de la qualité de l'enseignement obligatoire aux politiques du travail et de l'emploi, la DEPP est rendue incapable d'intervenir sur un terrain similaire sur fait de la construction de l'administration et par le fonctionnement en tuyau d'orgue des missions des directions ministérielles. Au-delà des études portant sur la politique économique, dont la mission revient logiquement au Ministère de l'économie et des finances, les évaluations portant sur la littératie des adultes par exemple, menées par l'OCDE et en lien direct avec les enjeux portés par PISA, sont suivies par le Ministère du travail. Celles sur la formation professionnelle également. Les évaluations portant sur les politiques d'enseignement supérieur sont suivies quant à elles par le Ministère ou secrétariat d'Etat de l'enseignement supérieur et de la recherche. Difficile pour la DEPP de pouvoir organiser une réponse unique aux implications de politique publique plus larges que seulement celles en lien avec la politique éducative et l'enseignement obligatoire.

La stratégie de valorisation des travaux de la DEPP par rapport à PISA est également mise à mal par de nouveaux enjeux internationaux récemment pris en compte par les pouvoirs publics, qui ont un impact direct sur l'imprévisibilité des besoins en compétences pour répondre aux besoins du marché du travail dans un avenir finalement proche : le réchauffement climatique, et les mouvements de population qu'il risque d'engendrer et qui sont difficiles à prévoir. Le constat de l'internationalisation, au sens d'homogénéisation des problèmes sociaux, a été posé dans le sillage des initiatives européennes de construction d'objectifs communs. Le réchauffement climatique, dont les effets commencent à se manifester de façon de plus en plus forte et régulière, rend la prise de conscience de l'immédiateté des enjeux encore plus prégnante, dans les discours des représentants de l'OCDE comme des décideurs politiques français. Eric

Charbonnier était ainsi par exemple convié en août 2018 à intervenir lors d'un colloque portant sur les grands défis de l'École dans une société mondialisée. Premier à prendre la parole, ses propos ont vite abordé l'imprévisibilité de l'avenir économique mondial et l'inadaptation qui en découle de l'évaluation des résultats de l'éducation à partir de programmes scolaires conçus avec les yeux d'aujourd'hui :

« Le monde évolue très rapidement et aujourd'hui nous sommes face à des enjeux que notre génération n'arrive pas à résoudre. On a le sujet du réchauffement climatique, l'émigration. On constate que beaucoup de métiers vont évoluer mais on ne sait pas exactement encore lesquels. On vit dans un monde qui évolue très vite et qui est comme on dit volatile incertain complexe et ambigu. Et ça veut dire qu'il faut préparer ces jeunes à ce monde de demain. Ça veut dire que le simple système d'éducation comme on le conçoit aujourd'hui, basé sur des connaissances, sur un individualisme, c'est ce qu'on voit dans le système français, finalement c'est quelque chose qui progressivement va disparaître pour faire place aux compétences, aux valeurs. C'est ce qu'on peut voir nous dans les pays de l'OCDE, de plus en plus de pays vont dans cette direction. »⁵⁰⁷

Ces propos sont tout à fait dans la ligne de ceux émis par Yves Durand en entretien sur l'homogénéité des problèmes à surmonter par les systèmes éducatifs européens, et qu'il a défendu au sein de la commission des affaires culturelles et de l'éducation de l'Assemblée nationale de 2009 jusqu'à la fin de son dernier mandat en 2017.

« Même si la politique éducative est du domaine national, elle se heurte à des sociétés qui évoluent un peu de la même manière et qui ont en gros les mêmes problèmes, et de plus en plus. L'hétérogénéité des populations est de plus en plus grande, on a des gens qui ne parlent pas la même langue, des différences sociales. On trouve ça dans tous les pays d'Europe, tous les pays de l'OCDE. Donc, si les systèmes ne sont pas les mêmes, les problèmes eux sont les mêmes. Les défis qui s'élèvent devant les systèmes scolaires sont les mêmes. D'où les réponses apportées par les uns et par les autres et qui peuvent être une inspiration. L'OCDE joue ce rôle-là. Et en même temps du

⁵⁰⁷ Colloque du 28 août 2018, Les grands défis de l'École dans une société mondialisée : de quelle École avons-nous besoin ? Entretiens Enseignants-Entreprises 2018 - Les entreprises dans la mondialisation.

même coup, elle est aussi une confirmation de la pertinence de ce qu'on fait ou pas, par rapport aussi à la réussite et aux échecs des autres. »⁵⁰⁸

Alors que la DEPP s'est efforcée de déconnecter PISA de son cadre international en en faisant une évaluation significative tant qu'elle est considérée comme nationale, le discours des décideurs politiques tend au contraire à internationaliser toujours plus les enjeux d'éducation nationale. PISA donne à voir non seulement l'état présent des compétences disponibles dans le pays, mais aussi sa compétitivité à venir face à des enjeux mouvants et des secteurs d'activité encore inconnus. La DEPP, avec son évaluation CEDRE encore centrée sur des programmes adaptés au « monde connu », est bien en peine de fournir de telles perspectives analytiques pour évaluer la pertinence d'orientations de politique éducative.

La mise en commun des enjeux interconnectés du fonctionnement du système éducatif, de l'enseignement supérieur, du marché du travail, des politiques de l'emploi et même des politiques sociales si l'on considère de travail de l'OCDE sur l'importance de lutter contre les inégalités scolaires, rend difficilement audible la stratégie de la DEPP de limiter l'intérêt des comparaisons spatiales entre pays de l'OCDE pour valoriser sa propre expertise.

3) L'enjeu de l'indépendance des évaluations et donc de leurs évaluateurs

Malgré les stratégies qu'elle déploie pour légitimer ses travaux auprès des décideurs publics, la DEPP peine donc à transmettre ses productions à ceux en charge du contrôle permanent de l'action de l'Etat, les parlementaires. Ses évaluations sont bien moins mobilisées par ces derniers que les publications des résultats PISA. Loin de s'éloigner au fur et à mesure des comparaisons

⁵⁰⁸ Entretien avec Yves Durand, Sénateur socialiste puis LREM de Mayenne, rapporteur de la loi refondation de l'école, président du comité de suivi et membre du comité d'évaluation des politiques publique, 14 juin 2017.

internationales peu pertinentes selon la DEPP pour se recentrer sur des comparaisons des résultats nationaux dans le temps, le « ranking » et le « benchmark » permis par la construction des indicateurs PISA restent des arguments forts employés par les parlementaires de tous partis politiques, et par le gouvernement en séance. On l'a vu, plusieurs raisons complémentaires expliquent ce défaut d'expertise attribué à la DEPP pour le contrôle continu de l'action de l'Etat en matière de politique éducative. Les travaux de la direction sont souvent jugés trop descriptifs et peu mobilisables immédiatement pour contribuer au débat sur les orientations souhaitables de la politique éducative française, et ses évaluations apportant une analyse des problèmes publics sont peu connues des parlementaires et bien moins mis en avant que l'évaluation PISA lorsqu'il s'agit de sujets sur lesquels toutes apportent un éclairage. Les travaux de la DEPP ne bénéficient pas de la même légitimité que PISA, qui bénéficie par ailleurs d'un contexte d'internationalisation des politiques éducatives qui place la comparaison des résultats des élèves au centre des préoccupations des décideurs politiques, et relègue les évaluations nationales à la fonction de confirmation plus que de révélation de problèmes publics.

Cette omniprésence de l'évaluation PISA dans les interventions au parlement, en particulier après la réforme constitutionnelle de 2008, ouvre toutefois un chemin pour les évaluations de la DEPP, qui se trouvent davantage mobilisées dans le sillage de PISA. Cette intensification de la légitimité accordée aux évaluations internationales et nationales est néanmoins porteuse d'un effet collatéral d'importance pour la DEPP. Alors que les parlementaires s'approprient progressivement l'outil argumentatif qu'est l'évaluation statistique, la multiplication des besoins de données pour asseoir leurs évaluations de l'action du gouvernement jointe à l'intensification du programme de travail de la DEPP fait renaître les interrogations sur l'indépendance de la direction du ministère, et donc sur la valeur d'expertise de ses productions pour évaluer l'action de l'Etat.

a) Evaluations externes ou internes : quel meilleur outil pour étayer le nouveau rôle d'évaluateur de l'action de l'Etat du Parlement ?

L'intensification des travaux d'évaluation parlementaire de l'action du gouvernement et de l'application des lois après la réforme constitutionnelle de 2008 s'accompagne d'un processus d'apprentissage de l'évaluation par les parlementaires. Les deux chambres décident d'incarner la réforme en mettant en place chacune de son côté plusieurs dispositifs pour remplir plus efficacement et plus clairement cette nouvelle mission officielle. Le comité d'évaluation et de contrôle des politiques publiques est mis en place à l'Assemblée nationale. Ce comité est conçu en transversalité des partis politiques et des commissions parlementaires. Il doit non seulement évaluer les politiques publiques grâce à un pouvoir d'auto-saisine sur les sujets qu'il considère devoir faire l'objet d'évaluation, mais aussi évaluer la qualité des études d'impact fournies par le gouvernement au moment du dépôt d'un projet de loi, et coordonner l'ensemble des travaux d'évaluation et notamment les rapports des missions d'information. Le Sénat quant à lui a créé deux instances permanentes, les délégations sénatoriales extra-législatives permettant d'informer le Sénat sur plusieurs thématiques larges comme la prospective ou les collectivités territoriales, et la commission sénatoriale pour le contrôle de l'application des lois chargée d'évaluer l'application des lois. L'évaluation est ainsi institutionnalisée clairement dans le fonctionnement des deux assemblées. Les différentes études menées sur l'appropriation de la démarche d'évaluation par les parlementaires et les matériaux recueillis dans le cadre de ce présent travail de thèse⁵⁰⁹ montrent d'abord deux choses : l'institutionnalisation de l'évaluation s'est accompagnée d'une volonté prudente des parlementaires d'occuper leur rôle d'évaluateurs de l'action publique, mais elle n'a pas entraîné de réelle montée en compétence de leur part sur la conceptualisation et les méthodes de l'évaluation de cette

⁵⁰⁹ Entretiens avec des parlementaires et collaborateurs parlementaires, étude des interventions en séance et lecture des rapports d'information en lien avec l'évaluation des politiques éducatives.

action⁵¹⁰, et donc sur les modalités d'utilisation optimales de toutes les évaluations à leur disposition – notamment celles de la DEPP.

A partir de la fin des années 2010, les usages des évaluations s'accroissent, et on voit évoluer dans le temps les habitudes des parlementaires d'aller chercher des données, de lire les rapports des évaluateurs externes et internes, pour construire une argumentation. Plusieurs travaux de recherche ont souligné le recours plus régulier, dans tous les secteurs de politique publique, aux travaux de la Cour des comptes ou du CESE, voire plus récemment de laboratoires universitaires⁵¹¹, à travers notamment la commande d'études et de rapports sur des sujets précis comme la prime d'activité ou encore l'impact des politiques d'exemption fiscales et sociales pour le développement des services à la personne⁵¹². Plus spécifiquement en matière de politique éducative et au-delà des études ayant un intérêt principalement budgétaire, les résultats de PISA – comme ceux de la DEPP d'ailleurs - ne sont pas uniquement ni même principalement utilisés par les parlementaires et les ministres qui leur répondent comme des outils de rationalisation de l'action publique⁵¹³ en éducation. L'approche partisane qui caractérise les groupes politiques au parlement freine ici toute volonté d'objectivisation des problèmes publics. Néanmoins les résultats PISA contribuent à intégrer l'usage de la statistique et de la démarche d'évaluation dans le quotidien du contrôle des politiques publiques. L'évolution dans le temps des fréquences d'utilisation des résultats des cycles PISA et des indicateurs produits dans le cadre du programme de l'OCDE montre bien l'accélération forte des

⁵¹⁰ Hortense de Padirac, « Le Parlement français et l'évaluation : une institutionnalisation impossible ? », *Sciences Po LIEPP*, working paper n°80, octobre 2018 ; Hortense de Padirac et Olivier Rozenberg, « L'évaluation au Parlement français : l'heure des choix », *Sciences Po LIEPP Policy Brief*, 2019.

⁵¹¹ Des études ont notamment été commandées à l'Institut des Politiques publiques (IPP), l'Observatoire français des conjonctures économiques (OFCE), le Groupement de recherche en économie quantitative (GREQAM) et le Laboratoire interdisciplinaire d'évaluation des politiques publiques (LIEPP) de Sciences Po. Voir la publication du LIEPP à ce sujet, Hortense de Padirac et Olivier Rozenberg, « L'évaluation au Parlement français : l'heure des choix. » *Sciences Po, LIEPP Policy Brief*, 2019.

⁵¹² Hortense de Padirac et Olivier Rozenberg, *art. cit.*

⁵¹³ Christian Maroy et Catherine Mangez, « Rationalisation de l'action publique ou politisation de la connaissance ? », *Revue française de pédagogie*, n°164, 2008.

références faites par les parlementaires aux résultats d'évaluation, que ces parlementaires soient d'ailleurs du même côté de l'échiquier politique que le ministre en place ou non. Si l'on compare les usages avant le dernier trimestre 2008 (soit après l'adoption de la réforme constitutionnelle) et après, les chiffres témoignent bien de l'augmentation des références aux résultats de PISA comme de CEDRE. Les mentions sont détaillées dans le tableau ci-dessous.

Tableau 1. Nombre de mentions de PISA et CEDRE dans les questions-réponses parlementaires à l'Assemblée Nationale et au Sénat, avant et après la réforme constitutionnelle de 2008.

		Assemblée Nationale	Sénat
Mentions de PISA	Avant la réforme constitutionnelle	22	2
	Après la réforme constitutionnelle	157	38
Mentions de CEDRE	Avant la réforme constitutionnelle	0	0
	Après la réforme constitutionnelle	34	2

La réforme constitutionnelle est, du point-de-vue chronologique, quasiment à la médiane de notre période d'étude. La remise de l'évaluation au cœur du débat sur le rôle des deux chambres parlementaires qu'elle a induite, accompagnée de la répétition des cycles PISA et CEDRE au fur et à mesure du temps, a contribué à banaliser l'usage des données que les deux évaluations fournissent. Elle a permis par ailleurs à PISA, et à son créateur l'OCDE, de se voir accorder un statut d'incontournable, publication après publication, que ce soit pour sa lecture des résultats généraux des élèves ou pour son analyse de problématiques clés qui trouvent une oreille particulièrement attentive au pays

de l'École républicaine, comme les inégalités scolaires⁵¹⁴. Les résultats et l'évaluation PISA sont sept fois plus mentionnés entre 2008 et 2019 qu'entre 2000 et 2008 à l'Assemblée nationale, et passent de zéro à plus d'une trentaine au Sénat. CEDRE suit la même tendance, même si les mentions sont bien moins nombreuses, en particulier au Sénat où elle n'est jamais utilisée à des fins argumentative ou politique avant 2016. La référence croissante de l'évaluation PISA au court des cycles est d'ailleurs tout à fait notable si l'on regarde le nombre de référence aux résultats publiés par l'OCDE entre chaque sortie de nouveau volume principal, tous les trois ans.

Tableau 2. Nombre de mentions de PISA dans les questions-réponses parlementaires entre la publication des résultats du 1^{er} cycle et la publication du 7^e cycle, à l'Assemblée nationale et au Sénat.

	Assemblée Nationale	Sénat
Entre les publications relatives au 1er et au 2e cycle (10 décembre 2001 - 9 décembre 2004)	4	0
Entre les publications relatives au 2e et au 3e cycle (10 décembre 2004 - 9 décembre 2007)	12	0
Entre les publications relatives au 3e et au 4e cycle (10 décembre 2007 - 6 décembre 2010)	12	6
Entre les publications relatives au 4e et au 5e cycle (7 décembre 2010 - 2 décembre 2013)	70	14
Entre les publications relatives au 5e et au 6e cycle (3 décembre 2013 - 6 décembre 2016)	68	10
Entre les publications relatives au 6e et au 7e cycle (7 décembre 2016 - 3 décembre 2019 ⁵¹⁵)	13	10

⁵¹⁴ Nous y viendrons au chapitre 6.

⁵¹⁵ Concernant les questions relatives aux résultats de PISA 2018, publiés donc en décembre 2019, il faut noter que bon nombre d'entre elles ont été posées en janvier 2020. Notre analyse s'arrêtant en 2019 pour permettre l'analyse des données et l'écriture de cette thèse, un certain nombre de mentions de PISA n'ont donc pas été comptées dans la dernière ligne de ce tableau. En tout état de cause, le contexte de grève au long cours au mois de décembre 2019 et janvier 2020 et le début de la pandémie de covid-19 en février 2020 ont contribué peut-être à réduire le nombre de questions posées sur la sortie des données PISA 2018.

Le gain d'intérêt progressif pour les données PISA à partir de la fin des années 2000 s'explique en partie par l'absence de données équivalentes suffisamment bien connues par les parlementaires produites par les évaluateurs français. Les statistiques jugées trop descriptives de la DEPP, bien identifiées car souvent citées en réponse aux questions par les ministres successifs, ne permettent pas de franchir le pas qui mène à l'analyse des problématiques lisibles dans les chiffres. Françoise Cartron, Sénatrice de Gironde depuis le 1^{er} octobre 2008, Vice-présidente du Sénat et de la commission de la culture, de l'éducation et de la communication, expliquait ainsi en entretien comment PISA était venu combler un manque dans les outils à la disposition des parlementaires pour mener à bien leur mission d'évaluation de l'action du gouvernement.

« Le rôle d'un sénateur, c'est à la fois de voter des lois, mais aussi de faire le suivi de leur application et de leur mise en œuvre. En particulier pour les politiques éducatives et pour tout ce qu'on a pu voter comme mesures dans le domaine éducatif. Il est évident que l'intérêt des choix que l'on fait, c'est de voir s'ils sont efficaces. Après, par rapport aux objectifs qu'on a fixés, on n'a pas d'outils au Sénat pour proprement évaluer cela. Je crois que quand je suis arrivé déjà il y a maintenant bientôt 8 ans, tout de suite on s'est tourné vers l'OCDE. Et en particulier M. Charbonnier, qu'on a auditionné puisqu'on se disait voilà, on va avoir un regard international, c'est-à-dire qu'on va dépasser notre propre regard purement hexagonal. On avait des indicateurs bien sûr, tant d'enfants qui ne savent pas bien lire, qui ne maîtrisent pas ci ou ça. En soi, c'est déjà quelque chose par rapport à l'objectif qu'on fixe. Mais je veux dire, il faut savoir aussi comment on se situe, et comment on se situe aussi internationalement, et ce que ça implique comme analyse des problèmes qu'on a. »⁵¹⁶

Les données et indicateurs de PISA viennent apporter un éclairage jugé indispensable à une évaluation complète en termes d'efficacité de la politique éducative française, et va au-delà de données descriptives qui ne fournissent pas aux parlementaires les clés d'analyse immédiate qu'ils attendent. On s'aperçoit

⁵¹⁶ Entretien avec Françoise Cartron, Sénatrice socialiste de Gironde, Rapporteuse de la loi refondation de l'école, 10 mai 2017.

d'ailleurs à la lecture des questions-réponses que les parlementaires ne questionnent plus les ministres à partir de la fin des années 2000 sur des chiffres sur l'état du système éducatif français. Alors qu'en début de période ils demandaient volontiers au ministre de leur fournir toute une série de chiffres sur l'état de l'école (par exemple le nombre d'élèves par niveaux, le nombre d'enseignants ou encore la prévision d'effectifs dans les écoles pour la rentrée à venir), ce type d'interrogation ne fait plus l'objet d'aucune question au gouvernement dans les années 2010. Cela ne signifie pas pour autant que les parlementaires renoncent à obtenir les chiffres bien évidemment, mais soit ils vont les chercher par eux-mêmes et emploient leurs questions à interroger le gouvernement sur leur signification pour l'état du système éducatif français, soit ils emploient des indicateurs qui problématisent directement les données descriptives pour en tirer des conclusions analytiques, comme peut le faire PISA. L'enjeu pour les parlementaires est bien ici d'aller à l'essentiel dans un programme de travail souvent fourni et éclaté entre plusieurs commissions, délégations et groupes de travail, et alors qu'ils sont souvent peu habitués aux méthodes d'évaluation et à l'analyse statistique.

Par ricochet, l'entrée dans le domaine du connu de la justification des propositions d'orientation de politique éducative par les résultats de PISA contribue à faciliter les références aux travaux de la DEPP, qui sont également de plus en plus repris par les parlementaires dans leurs questionnements et par les ministres dans leurs réponses. Toutes les évaluations tirent profit de l'aura de PISA, et la DEPP bénéficie de cette popularisation de l'usage de la statistique pour étayer l'argumentation. CEDRE se voit progressivement convoqué soit seul, soit le plus souvent aux côtés de PISA et sur un même pied d'égalité. Cette reprise des conclusions de CEDRE et leur mise en parallèle régulière avec PISA sont particulièrement visibles dans les questions au gouvernement et dans les rapports à partir de la fin de l'année 2011. Entre décembre 2011 et décembre 2014 en particulier, l'évaluation CEDRE sera citée dans 34 échanges lors de questions parlementaires, dont 23 fois par le ministre dans sa réponse. Dans 24 de ces échanges, CEDRE est nommé comme le binôme de PISA, permettant d'avoir

une vision complète, nationale et internationale, des résultats des élèves français, et porteur d'un niveau de preuve similaire. De la même façon, les références à CEDRE dans les rapports parlementaires s'accroissent après 2010, et sont toujours accompagnées, ou plutôt accompagnent toujours, les résultats de PISA. Les résultats de CEDRE apparaissent notamment régulièrement dans les rapports préalables à l'examen des lois de finances annuelles, pour contextualiser les jugements portés par les rapporteurs sur les choix budgétaires du gouvernement pour l'année à venir. Ils apparaissent également après 2013 à l'Assemblée nationale et après 2015 au Sénat dans quelques rapports et une proposition de résolution⁵¹⁷. Les tendances de résultat étant la plupart du temps proches, les données CEDRE servent à confirmer ce que dit PISA à des échéances plus resserrées et avec un découpage par discipline plus familier aux décideurs politiques français, puisque le cycle d'évaluation de la DEPP se calque sur les disciplines telles que présentées dans les programmes scolaires. Si PISA se voit toujours accordé une place de premier rang dans les questions et les rapports parlementaires au détriment des évaluations faites par la DEPP, il faut néanmoins noter que les éléments de langage de CEDRE sont ceux autour desquels sont centrées régulièrement les analyses formulées. On retrouve ainsi plus régulièrement mention de constats d'une baisse de niveau ou d'une chute dans le classement en mathématiques, français, histoire-géographie, qu'en « maîtrise de la langue » ou « compétences en sciences », pourtant objets d'évaluation issus de PISA et porteurs d'un sens autre que les dénominations des disciplines des programmes scolaires français⁵¹⁸.

⁵¹⁷ Au Sénat, une proposition de résolution sur la création d'une commission d'enquête sur le fonctionnement du service public de l'éducation, et un rapport d'information en 2015 sur le thème « *Faire revenir la République à l'École* », donnant les résultats de la commission d'enquête précitée. A l'Assemblée nationale, un rapport d'information du comité d'évaluation et de contrôle des politiques publiques sur « *L'évaluation des politiques publiques en faveur de la mixité sociale dans l'éducation nationale* » et un rapport d'information du même comité, de Régis Juanico et Jean Frédéric Poisson sur « *L'évaluation des politiques publiques en faveur de la mobilité sociale des jeunes* ». CEDRE est également cité dans le rapport sur le projet de loi d'orientation et de programmation pour la refondation de l'école de la république, 28 février 2013, présenté par Yves Durand.

⁵¹⁸ Voir le chapitre 2 sur la construction de PISA pour plus de détails.

L'habitude est prise après 2008 d'étayer ses évaluations et de formuler ses recommandations à partir de chiffres externes aux deux chambres. Elles tendent soit à déléguer aux instances évaluatives qu'elles jugent légitimes, soit à piocher dans les évaluations existantes également légitimes. PISA et CEDRE se trouvent dans la deuxième catégorie à la fin de notre période, conférant à ceux qui les produisent un rôle de facto d'expert de l'action publique en éducation. Ni l'Assemblée nationale ni le Sénat ne se sont lancés dans la production de leurs propres évaluations, malgré des velléités restées inabouties de certains parlementaires. Plusieurs députés lancent notamment en 2018 l'idée de créer une cellule d'évaluation budgétaire à l'Assemblée nationale qui soit composée d'une équipe indépendante d'économistes et de statisticiens sur un modèle anglo-saxon. Une tribune de presse publiée le 18 avril 2018 dans *Le Monde*, intitulée « Pour un débat budgétaire responsable et libéré de l'arbitraire », signée par des parlementaires de tous groupes politiques, formule des propositions concrètes reprises par un groupe de travail sur les moyens de contrôle et d'évaluation du Parlement en juin de la même année. L'ambition de cette instance nouvelle est de fournir des données « indépendantes », c'est-à-dire obtenues sans que le ministère ne puisse éventuellement influencer leur objet (ce sur quoi elles portent), leur publication (en ne fournissant que les données allant dans le sens des réformes souhaitées et passant sous silence les autres moins favorables), ou même leur résultat (en ne fournissant pas les « vrais » chiffres). Le projet, abandonné au moment du changement de présidence de l'Assemblée nationale fin 2018, constitue néanmoins un moment d'apogée d'une tendance à la méfiance envers l'outil évaluation, jugé potentiellement partisan et peu objectif malgré des formes employées souvent louées pour leur rationalité, les indicateurs statistiques notamment.

b) L'indépendance de la DEPP au cœur des questionnements

Ce rôle désormais central de la statistique dans la naissance et la vie des problèmes publics contribue paradoxalement à entraver la construction par la

DEPP de son rôle d'expert de la politique éducative française. Ce paradoxe est nourri par la méconnaissance, sauf exception, par les parlementaires des théories statistiques et des principes qui sous-tendent la construction des évaluations nationales et internationales. Il n'est évidemment pas demandé aux parlementaires d'avoir des compétences particulières en statistiques, les administrations des assemblées parlementaires sont présentes pour pallier ce manque du point-de-vue technique, et les collaborateurs parlementaires ont théoriquement un rôle d'accompagnement pour assurer un usage politique juste des données qu'ils lisent dans les rapports à leur disposition ou dans la presse. Néanmoins, dans les faits, rares sont ceux qui saisissent pleinement les tenants et les aboutissants des évaluations qu'ils utilisent dans leurs argumentaires politiques et dans les travaux d'évaluation des politiques publiques dans le cadre de leur mandat. L'une des conséquences de cette lecture rapide des résultats des évaluations est la crainte chez certains d'une manipulation de ce nouvel emploi du chiffre⁵¹⁹, par la construction artificielle de statistiques qui seraient orientés pour rendre visibles certains problèmes et en masqueraient d'autres peut-être plus importants, plus urgents, aux répercussions plus larges et/ou plus délétères. Elle rend la direction interne au ministère régulièrement suspecte aux yeux de ceux qui doivent évaluer l'action de l'Etat et la bonne application des lois. En miroir, PISA est jugé être un programme indépendant des enjeux nationaux, mené par une organisation internationale aux intérêts propres. Jamais ne transparaît dans les usages des parlementaires l'intervention pourtant revendiquée de la DEPP dans la construction et la mise en œuvre de PISA pour la France.

Les évaluations et indicateurs produits par la DEPP, qui peuvent donner lieu aussi bien à une description chiffrée du système éducatif (nombre d'élèves par niveau présents, attendus, taux d'absentéismes constatés, nombre d'enseignants par niveau, heures d'enseignement par discipline, ...) qu'à des questionnements problématisés pouvant donner lieu à des réorientations de politiques éducatives (l'impact du redoublement sur les résultats des élèves, la

⁵¹⁹ Erik Neveu, « La construction sociale des problèmes publics », *Idées économiques et sociales*, n°190, décembre 2017, p.14.

maîtrise de la langue à différents niveaux ou selon les types d'établissements, ...) sont principalement convoqués par les ministres dans les réponses faites aux questions des parlementaires. L'analyse de l'emploi des données ou des conclusions d'évaluation par les parlementaires et par les ministres montre bien cette habitude d'usage par les ministres des données de la DEPP, bien plus que par les parlementaires. Les références à PISA sont au contraire quant à elles largement faites par les parlementaires. Au Sénat, 75% des références à PISA se trouvent dans les questions posées par les sénateurs, et 25% sont faites par les ministres dans leurs réponses. A l'Assemblée Nationale, la proportion de références faites par les députés est de 66%, et 34% par les ministres. Si l'on regarde la même répartition des références pour les travaux de la DEPP, elle est inverse : 70% des références aux travaux de la DEPP au Sénat sont faites par les ministres, 76% à l'Assemblée Nationale.

Les ministres successifs ou leurs représentants interrogés sur la réaction du ministère à tel ou tel résultat d'évaluation répondent quasi-systématiquement en utilisant les données produites par la DEPP. Rien d'anormal bien sûr, puisque la DEPP a entre autres pour mission de fournir les données relatives à l'état du système éducatif français et doit répondre aux demandes de précisions du ministre lorsque des dimensions descriptives lui manquent pour argumenter ses prises de position ou ses décisions. Mais cela a pour corollaire d'associer la DEPP à son ministère de tutelle par ceux qui reçoivent les réponses des ministres. Cela transparait notamment dans la façon dont les parlementaires se réfèrent à la DEPP. Ils la décrivent régulièrement comme l'organe produisant les « chiffres du ministère », « les statistiques du ministère », et les données produites sont donc directement associées aux résultats de l'action du ministère, et par dérivation du ministre.

« Selon une étude menée par la direction de l'évaluation et de la prospective de son ministère à la fin de collège, si un élève sur quatre a des performances manifestant une maîtrise complète des compétences

générales attendues, 15 % font preuve d'une maîtrise réduite de ces compétences. »⁵²⁰

« Ainsi, les statistiques des services de votre ministère indiquent que 79 % des élèves provenant de catégories sociales favorisées obtiennent un bac général, contre seulement 18 % des élèves issus de milieux défavorisés. »⁵²¹

La suspicion de l'existence d'un lien de subordination entre la DEPP et le ministre de l'Education nationale n'est pas nouvelle. Elle a déjà été formulée à de nombreuses reprises, y compris par un ministre de l'Education nationale, Claude Allègre, au début des années 2000⁵²², qui avait initié un mouvement de déshabillage de la direction qu'il jugeait tout à fait illégitime pour évaluer le système éducatif⁵²³.

Il faut dire que les réponses des ministres aux parlementaires tendent à plusieurs reprises à entretenir le flou sur les frontières entre l'action politique du ministère et celle de son administration évaluatrice. Elles témoignent de conceptions très personnelles et très différentes les unes des autres du rôle de la DEPP, qu'ils énoncent parfois clairement. Certains, comme Najat Vallaud-Belkacem, confirment fermement l'indépendance de la DEPP, notamment lorsqu'elle produit des indices sociaux pour analyser les résultats de ses évaluations, indices aux implications potentiellement explosives en termes de critique des résultats de la politique éducative menée jusque-là.

« Afin d'élaborer la nouvelle carte de l'éducation prioritaire, dont les contours n'avaient pas été revus depuis 2006, un indice social innovant a été défini, en toute indépendance, par la direction de l'évaluation, de la prospective et de la performance du ministère afin de prendre davantage en compte les critères sociaux, notamment la situation des familles. »⁵²⁴

⁵²⁰ Question écrite n°56726 de Monsieur Christian Estrosi posée à Monsieur le ministre de l'Education nationale publiée au JO de l'Assemblée Nationale le 1^{er} janvier 2005.

⁵²¹ Question crible thématique (question orale) n°0156C de Madame Jacqueline Gourault, posée à Monsieur le ministre de l'Education nationale publiée au JO du Sénat le 9 février 2011.

⁵²² Lorraine Data, *Le grand trucage. Comment le gouvernement manipule les statistiques*, Paris, La Découverte, 2009.

⁵²³ Voir chapitre 1 pour plus de détails.

⁵²⁴ Réponse du ministre à la question orale n°0974S de Madame Aline Archimbaud, Seine-Saint-Denis, écologiste, publiée au JO du Sénat le 25 décembre 2014.

D'autres à l'inverse contribuent plutôt, volontairement ou non, à entretenir la confusion des rôles, en présentant leur relation de travail avec la DEPP de façon verticale, où leur position est celle d'un commanditaire. La réponse de Jean-Michel Blanquer à la sénatrice de Gironde Nathalie Delattre, qui l'interrogeait sur la façon dont les zones d'éducation prioritaire allaient être redéfinies pour éventuellement prendre en compte la ruralité des territoires, est très représentative de ce positionnement ambigu :

« Vous avez raison, la définition des zones d'éducation prioritaires repose peut-être, jusqu'à présent, sur des critères qui ne permettent pas de prendre en compte l'ensemble des élèves défavorisés de France. Nous devons donc la faire évoluer pour que le dispositif soit plus dynamique, plus efficace et plus juste.

J'ai demandé à la direction de l'évaluation du ministère de calculer l'indice d'éloignement, dont vous avez souligné l'intérêt. Les premiers résultats sont parvenus hier : nous avons donc maintenant techniquement les moyens d'intégrer ce critère, qui est en effet très important. »⁵²⁵

Si cela semble bien entrer dans les attributions officielles de la DEPP, la formulation du ministre laisse sous-entendre un rapport de subordination entre la DEPP et la personne du ministre de l'Éducation nationale. La commande, suivie de la rapidité d'exécution, induit un rapport hiérarchique n'allant pas dans le sens de la réaffirmation de l'indépendance du programme de travail de la direction.

Les risques perçus par les parlementaires de collusion entre le ministre et « sa » direction de l'évaluation alimentent, au début des années 2010, les craintes de certains d'entre eux sur le respect effectif de certains principes clés de l'indépendance de l'évaluation. C'est notamment le cas du principe de respect des personnes⁵²⁶, impliquant l'anonymat des répondants et le respect du principe

⁵²⁵ Question au gouvernement n°1017G de la sénatrice Nathalie Delattre, Gironde, RDSE, publiée au JO le 7 novembre 2019.

⁵²⁶ Société Française de l'Évaluation, *Charte de l'évaluation des politiques publiques et des programmes publics*, actualisée le 21 juin 2006. La Charte de la SFE est communément acceptée par les acteurs de l'évaluation en France comme présentant les principes déontologiques nécessaires à

de protection des données personnelles selon les règles de la CNIL. Plusieurs parlementaires au fil des ans se font les porte-paroles de la crainte de voir l'identité des élèves et des professeurs utilisée d'abord dans les évaluations de la DEPP pour être ensuite fournie librement aux chefs d'établissement ou au ministère pour être employée à d'autres fins. Le sujet est d'autant plus sensible que la DEPP s'engage à l'époque sur des recueils de données allant au-delà des résultats des élèves à des évaluations disciplinaires, avec par exemple des études sur l'engagement syndical des enseignants ou la constitution d'évaluation sur les conditions de scolarité des élèves.

« Mme Danièle Pourtaud attire l'attention de M. le ministre de la jeunesse, de l'éducation nationale et de la recherche sur une enquête de " fin de scolarité au collège " diligentée par la direction de l'évaluation et de la prospective du ministère de la jeunesse, de l'éducation et de la recherche, parvenue dans certains établissements scolaires au cours du mois de mai. Cette enquête, certes déclarée à la CNIL (Commission nationale de l'informatique et des libertés), fait apparaître tant dans la forme que sur le fond plusieurs problèmes éthiques.

(...)Le corps enseignant s'inquiète de l'élaboration de tels documents qui créent des profils et des analyses de comportements mis à la disposition des chefs d'établissement et des inspections académiques, au-delà des procédures de contrôle normales dans le cadre du suivi pédagogique et de carrière des enseignants. (...)Mme Pourtaud s'étonne que l'administration de l'éducation nationale classe ainsi professeurs et élèves dans un fichier d'enquête dont la confidentialité reste toute relative, les élèves pouvant comparer leurs réponses respectives. Elle lui demande si, au-delà de la simple déclaration à la CNIL, il n'aurait pas été préférable de soumettre à cette instance le contenu du questionnaire pour délibération.

«Réponse du ministère : (...) L'ensemble de cette évaluation est inscrit au programme officiel des opérations statistiques organisées par la direction de l'évaluation et de la prospective, publié au BOEN n° 1 spécial du 23 janvier 2003. À ce titre, elle est couverte par le secret statistique. »⁵²⁷

la conduite d'évaluations de l'action publique de qualité. Ils sont au nombre de sept : pluralité, distanciation, compétence, respect des personnes, transparence, opportunité et responsabilité.

⁵²⁷ Question écrite n°08098 de Madame Danièle Pourtaud, Paris, Socialiste, publiée au JO du Sénat le 19 juin 2003 et réponse du Ministère de la jeunesse, de l'éducation nationale et de la recherche publiée au JO du Sénat le 7 août 2003.

La même crainte de voir les données sur les élèves et les enseignants reprises de façon partisane est formulée par le sénateur Gaëtan Gorce en 2012, à propos d'une évaluation sur l'absentéisme, laissant craindre la constitution de fichiers individuels sur les élèves absentéistes fournis aux rectorats, plutôt qu'une étude longitudinale anonyme sur l'évolution du taux d'absentéisme général dans les établissements scolaires français.

« M. Gaëtan Gorce interroge M. le ministre de l'éducation nationale sur la loi n° 2010-1127 du 28 septembre 2010 visant à lutter contre l'absentéisme scolaire qui a modifié l'article L. 131-8 du code de l'éducation nationale. (...) Il paraît indispensable de clarifier le cadre applicable afin de prévenir le développement de traitements automatisés de données relatives au dispositif de suivi des élèves absentéistes. (...).

Réponse du ministère : L'article L. 131-8 du code de l'éducation nationale, dans sa rédaction issue de la loi n° 2010-1127 du 28 septembre 2010 visant à lutter contre l'absentéisme scolaire, dispose que le directeur académique des services de l'éducation nationale doit réaliser un suivi des élèves absentéistes. Cette préconisation n'a pas pour objet la constitution de fichiers individuels aux seuls fins de lister l'ensemble des élèves absentéistes. À ce jour, seuls les indicateurs élaborés par la direction de l'évaluation, de la prospective et de la performance du ministère de l'éducation nationale, sur la base d'enquêtes anonymes par échantillon, permettent un suivi de l'absentéisme. »⁵²⁸

La réponse officielle du ministère est de rassurer sur l'usage des données en valorisant l'expertise de la DEPP menant ses évaluations dans le respect des règles de protection des données personnelles et d'anonymat des réponses. Il rassure également sur sa pleine compétence pour mener des études sur échantillon, donc dans un but d'analyse et d'aide à l'orientation de la politique éducative, et non pour produire des données descriptives de situations individuelles.

Les prises de position sur la nature d'évaluateur indépendant de la DEPP, et par dérivation sur l'aide à la décision publique apportée par les données qu'elle

⁵²⁸ Question écrite n°03226 de Monsieur Gaëtan Gorce, Nièvre, Socialiste, publiée au JO du Sénat le 22 novembre 2012.

produit, sont mises en débat public officiellement lors de l'examen de la loi d'orientation et de programmation pour la refondation de l'école de la République en 2013. La loi propose de créer le CNESCO, Conseil national d'évaluation du système scolaire, dont les missions sont triples. Elles sont présentées officiellement, une fois la création du CNESCO actée dans la loi, dans la lettre de mission envoyée par le ministre Vincent Peillon à la Présidente du CNESCO Nathalie Mons. Le nouveau conseil est chargé de la « *production d'évaluations et de synthèses d'évaluations, notamment dans une perspective internationale, d'expertise méthodologique des évaluations existantes, et de promotion de la culture de l'évaluation en direction des professionnels de l'éducation et du grand public.* »⁵²⁹ L'OCDE et son évaluation PISA est citée nommément dans la lettre de mission, comme la révélatrice de l'effet particulièrement fort de l'appartenance sociale sur les résultats scolaires des élèves, et le ministre demande expressément au CNESCO de se saisir de ce « *problème grave* ». Une fois encore, PISA met à l'agenda un problème public pourtant identifié de longue date et travaillé par la DEPP dans plusieurs de ses évaluations. La direction du ministère n'est même pas mentionnée par le ministre, dont on suppose qu'il l'intègre dans « *l'ensemble des directions du ministère de l'Éducation nationale* » à qui il demande de mettre à la disposition du CNESCO « *les données statistiques nécessaires à la conduite de (ses) missions* ». Cette distribution des rôles et l'assimilation de la DEPP à une direction comme une autre du ministère, sous l'influence directe de ministre, a transparu tout au long des débats parlementaires portant sur la création du CNESCO. Le rapport produit par Yves Durand est particulièrement clair sur la position de l'Assemblée nationale. Il classe ainsi les instances d'évaluation principales utilisées dans le cadre du pilotage de la politique éducative française.

« Plus généralement, les instances d'évaluation de notre système éducatif sont, aujourd'hui, soit non indépendantes, soit « externalisées », soit mobilisées par d'autres missions, ce qui n'est évidemment guère satisfaisant :

⁵²⁹ Lettre de mission du 24 janvier 2014 de Vincent Peillon, ministre de l'Éducation nationale, à Nathalie Mons, Présidente du CNESCO.

– non indépendantes. En effet, le décret n° 2006-572 du 17 mai 2006 fixant l'organisation de l'administration centrale des ministères de l'éducation nationale et de l'enseignement supérieur et de la recherche confie à la direction de l'évaluation, de la prospective et de la performance (DEPP) la conception et la mise en œuvre, à la demande des autres directions, d'un programme d'évaluations, d'enquêtes et d'études sur tous les aspects du système éducatif et de recherche. Ce service doit assurer, par ailleurs, la cohérence de la mesure de la performance aux niveaux national et territorial et calculer les indicateurs de performance que les directions et les services déconcentrés mettent en œuvre dans les programmes budgétaires ;

– externalisées. À l'heure actuelle, ce sont, dans une large mesure, les évaluations de l'OCDE, c'est-à-dire PISA, et les enquêtes de la Cour des comptes sur l'éducation nationale (165) qui, aux yeux des médias et du public, permettent de diagnostiquer les performances de l'école. Peut-on cependant s'appuyer sur les seuls travaux, publiés tous les trois ans, d'une organisation internationale pour piloter notre système éducatif ? Cette question vaut aussi pour les rapports – remarquables – de la Cour des comptes, car cet organe de contrôle n'a pas pour seule mission d'assurer l'évaluation du ministère de l'éducation nationale.

La situation est d'autant plus paradoxale qu'il existe des indicateurs nationaux, peu connus, mais rigoureux pour mesurer les acquis des élèves au regard des objectifs fixés par les programmes disciplinaires : le cycle des évaluations disciplinaires réalisées sur échantillon (CEDRE), ce dispositif ayant été mis en place en 2003 ;

– mobilisées par d'autres missions. C'est le défaut du HCE qui exerce son « droit de regard » consultatif sur l'élaboration des programmes, sur les résultats du système éducatif et sur la formation des enseignants (...). »⁵³⁰

La DEPP y est présentée comme « non indépendante », alors même qu'elle met régulièrement en avant ses modalités de travail, notamment pour la conception de son programme annuel de travail, qui lui confèrent au moins partiellement une indépendance théorique. CEDRE est quant à lui encore une fois associée à

⁵³⁰ Haut conseil de l'évaluation de l'école, *La France et les évaluations internationales des acquis des élèves*, avis n° 16, mai 2005.

PISA, et présentée comme un challenger possible pour l'évaluation de l'OCDE dans le pilotage du système éducatif français. Son intérêt et sa rigueur méthodologique lui sont reconnus, l'ancienneté de ses premiers résultats est valorisée, la validité de l'objet qu'elle mesure – c'est-à-dire les résultats par rapport au programme scolaire – est bien perçue. L'opération manquante reste celle de la popularisation de ses résultats, regrettée par les auteurs du rapport parlementaire. Il est intéressant de noter ici que la DEPP n'est pas mentionnée comme la créatrice et la porteuse de CEDRE, peut-être afin de ne pas réduire la portée de l'intérêt formulé d'une étude à la main d'une instance « non-indépendante ».

La position de non-indépendance de la DEPP est partagée au Sénat. La sénatrice Françoise Cartron, rapporteure de la loi de 2013 au Sénat comme le fût Yves Durand à l'Assemblée nationale, l'exprimait en entretien.

« Il faut une évaluation indépendante, et jusque-là ceux qui faisaient ça c'était la Maison Education nationale en interne, ils faisaient leurs propres évaluations avec les inspections générales et la direction de l'évaluation. Ils n'ont pas suffisamment de liberté, il y a eu des évaluations à charge qui étaient mises au panier par le ministre. Et donc on a obtenu du ministre qu'il y ait la création de son Conseil national des évaluations, qui soit cette fois totalement indépendant. »⁵³¹

Difficile pour la DEPP de faire valoir l'indépendance de ses travaux d'évaluation et notamment de ses travaux prospectifs. Le rapport du comité d'évaluation et de contrôle des politiques publiques publié en septembre 2018 et portant sur « L'organisation de la fonction d'évaluation du système éducatif ⁵³² » positionne la DEPP comme un pourvoyeur de données au CNESCO, au même plan que les inspections générales, sans mentionner à aucun moment le choix fait indépendamment des commandes ministérielles de son programme annuel de travail.

⁵³¹ Entretien avec Françoise Cartron, Sénatrice socialiste de Gironde, Rapporteure de la loi refondation de l'école, 10 mai 2017.

⁵³² Rapport d'information du comité d'évaluation et de contrôle des politiques publiques de l'Assemblée nationale, portant sur l'organisation de la fonction d'évaluation du système éducatif, présenté par Régis Juanico et Marie Tamarelle-Verhaeghe, le 27 septembre 2018.

Le débat ne s'est pas clos avec la création du CNESCO, valorisant l'intérêt de PISA comme évaluation clé pour l'avenir du système éducatif français et allouant à la DEPP le rôle de direction dépendante des orientations voulues par le ministre, intéressante par les données descriptives qu'elle produit. Le CNESCO, dès ses premiers travaux, n'a pas convaincu les parlementaires de sa capacité à fournir des évaluations du système scolaire totalement indépendante, et notamment Yves Durand qui avait été rapporteur de la loi créant le Conseil national d'évaluation.

« Alors du coup, comme on n'avait aucun organe d'évaluation indépendant en France, on a créé dans la loi le CNESCO, le Conseil national d'évaluation du système scolaire, mais c'est un raté. C'est un raté pour plusieurs raisons. Il ne fait pas vraiment de l'évaluation. Ce n'est pas vraiment indépendant du ministère. Alors qu'un système d'évaluation, c'est indépendant. Il faut que l'évaluation soit indépendante. Moi, j'étais pour la création d'une agence indépendante comme il existe au Québec, comme l'OCDE le préconise. »⁵³³

L'OCDE contribue ici, dans l'exemple pris par Yves Durand, à construire et à diffuser une définition propre de l'indépendance d'une évaluation, qui ne va pas nécessairement dans le sens de celle soutenue par la DEPP forcément liée à son existence au sein du ministère de l'Éducation nationale.

Les mêmes interrogations sur l'indépendance des évaluations nationales du système éducatif français et du CNESCO resurgissent une nouvelle fois les derniers mois de notre période d'étude, lors de la loi pour une école de la confiance portée par Jean-Michel Blanquer et promulguée le 28 juillet 2019. Le texte proposé par le ministre prévoit en effet de supprimer le CNESCO pour en faire un Centre national d'étude des systèmes scolaires, devant désormais faire valider son programme de travail par le ministère, et ne pouvant plus produire ses propres évaluations. La DEPP conserve encore une fois son rôle de pourvoyeuse de données, sans réussir réellement à mener au bout sa stratégie de revalorisation de son expertise non seulement méthodologique – que tout le

⁵³³ Entretien avec Yves Durand, Sénateur socialiste puis LREM de Mayenne, rapporteur de la loi refondation de l'école, président du comité de suivi et membre du comité d'évaluation des politiques publiques, 14 juin 2017.

monde lui reconnaît – mais aussi analytique dirigée vers l'action publique en éducation.

Conclusion du chapitre

La stratégie mise en place par la DEPP lors de la publication des premiers résultats PISA avait pour objectif de positionner la direction parmi les évaluateurs ressources clés des décideurs politiques, et notamment dans le cadre du contrôle continu de l'action de l'Etat en matière de politique éducative. Pour cela, elle a mis en place un plan d'action en trois grands axes que nous avons détaillé au chapitre 4. D'une part, elle a souhaité travailler à la mise en place d'outils du même acabit que PISA, et a développé son évaluation CEDRE. Deuxième axe de son travail de positionnement parmi les experts de l'évaluation des politiques publiques, elle a œuvré pour expliciter la méthodologie et les conditions de validité de PISA, en remplaçant l'évaluation internationale dans un cadre national et en valorisant les comparaisons temporelles permises par le fonctionnement en cycles, plutôt que les comparaisons internationales. Enfin, son troisième axe de travail a été de renforcer son lien avec le monde de la recherche afin de rappeler définitivement le principe d'indépendance et de prospective régissant une partie de ses travaux.

L'étude de l'usage des travaux de la DEPP dans le processus de contrôle continu de l'action de l'Etat au Parlement montre bien que la stratégie de la direction du ministère porte difficilement ses fruits à la fin de notre période. Le positionnement d'entrepreneur de problèmes publics qu'elle cherche à avoir auprès des parlementaires n'est pas assuré par sa capacité à mettre à l'agenda les problèmes publics qu'elle peut arriver à justifier et à populariser, et cela même si ses travaux sont souvent repris à la fois par les sénateurs et députés et par les ministres successifs. Si les travaux de quantification des différentes masses du système éducatifs, utilisés par ailleurs pour certains pour intégrer les indicateurs

de pilotage budgétaire issus de la LOLF, sont parfois utilisés par les parlementaires pour appuyer l'argumentation de leur interpellation du gouvernement et du ministre de l'Education nationale en particulier, les évaluations menées dans le cadre de leur programme de travail annuel sont quant à elles peu reprises.

Ce n'est pas le cas de PISA qui est régulièrement employé comme argument venant appuyer la problématique mise en lumière dans le cadre des questions posées au gouvernement, ou comme fait représentatif d'une situation de crise de laquelle devrait découler une réaction du gouvernement. Les indicateurs et les conclusions publiés par l'OCDE se voient conférer une légitimité pouvant donner lieu sans réserve à une traduction des résultats en objet de politique éducative, d'un savoir scientifique objectif en politique publique. PISA supplante d'ailleurs toutes les autres évaluations dans les discours politiques des parlementaires comme des ministres. Les grands moments de communication organisés autour de la publication des résultats triennaux portent certes particulièrement jusqu'aux parlementaires qui se saisissent de façon coordonnée des indicateurs juste parus pour faire part de leurs interrogations et inquiétudes sur l'état du système éducatif français. Mais de façon générale, l'utilisation des données PISA est constante au fil des ans, et contribue à conférer à l'OCDE le rôle d'un pourvoyeur reconnu et légitime de problèmes publics devant être mis à l'agenda politique. Si l'évaluation internationale est davantage convoquée par les parlementaires et ministres de droite que ceux de groupes de gauche, surtout dans la première décennie de sa mise en œuvre, personne n'en critique ni le bienfondé ni la valeur méthodologique.

L'internationalisation croissante des enjeux économiques et d'emploi, matérialisés au début des années 2000 par la stratégie de Lisbonne puis la stratégie européenne pour l'emploi, donne par ailleurs à PISA un poids particulier lorsqu'il s'agit d'imaginer les politiques éducatives et de formation de demain permettant à la France d'occuper une place de choix parmi les pays les plus attractifs et les plus compétitifs. PISA, objet de discussion dans les réunions des décideurs politiques européens et internationaux, devient alors l'outil de

référence pour évaluer le positionnement de la France sur ces sujets. Les travaux de la DEPP non seulement ne permettent pas d'éclairer la place de la France parmi ses partenaires et concurrents puisqu'ils ont un cadre national, mais ils ne peuvent en plus pas servir d'outil utilisé par les décideurs politiques français dans les réunions internationales pour justifier de leurs résultats. La force de la comparaison internationale permise par PISA est particulièrement importante à la fin de notre période d'étude, et n'est en aucun cas supplantée par la force des comparaisons temporelles entre les cycles pour un même pays. La légitimité de l'OCDE auprès des parlementaires français pour remplir leur mission de contrôle continu de l'action de l'Etat et d'évaluation de l'action publique est donc considérablement renforcée par l'outil PISA.

La DEPP voit paradoxalement sa légitimité d'experte questionnée en même temps que ses évaluations sont de plus en plus utilisées dans les échanges entre parlementaires et gouvernement. Ses travaux sont régulièrement associés aux orientations politiques voulues par le ministre, et l'indépendance de la direction est au mieux questionnée ouvertement, au pire niée en considérant qu'elle obéit aux commandes ministérielles. La validité de l'emploi de ses données pour évaluer l'action du gouvernement en matière de politique éducative est donc réduite au minimum, c'est-à-dire à un usage des statistiques descriptives qu'elle produit, mais pas réellement des orientations qu'elle propose. Néanmoins, il faut tout de même noter que l'évaluation CEDRE devient plus régulièrement citée aux côtés de PISA, signe que la stratégie d'inspiration adoptée par la DEPP pour créer son évaluation maison a rencontré un relatif succès, et que ses travaux sont en partie utilisés à des fins d'apport de connaissance lors de débats portant sur le fonctionnement du système éducatif français.

Chapitre 6 : PISA et la réforme éducative : l'expertise internationale au cœur de la fabrique de la loi nationale

Ce sixième chapitre s'inscrit dans une démarche d'observation de l'évaluation statistique comme expertise utilisée aux moments critiques⁵³⁴ d'action publique en éducation, et plus spécifiquement aux grands moments réformateurs de la période 2000 – 2017. Nous étudierons ici deux réformes de fond du système éducatifs français portées par deux gouvernements aux couleurs politiques différentes : la loi de programmation et d'orientation de 2005⁵³⁵ portée principalement par François Fillon⁵³⁶, et la loi de programmation et d'orientation de 2013 portée par Vincent Peillon⁵³⁷.

Chronologie et grands principes de la loi de 2005 pour l'avenir de l'École et de 2013 pour la refondation de l'école de la République

La loi de 2005 pour l'avenir de l'École :

En 2004, Luc Ferry, alors ministre de l'Éducation nationale lance le débat national sur l'avenir de l'école. Le rapport de la concertation est remis au premier ministre le 10 octobre 2004. C'est un préalable au texte de loi dont le projet est déposé le 12 janvier 2005.

⁵³⁴ La terminologie est employée notamment par Pierre Clément qui en fait une rétrospective dans *Réformer les programmes pour changer l'école. Une sociologie historique du champ du pouvoir scolaire*. Thèse pour le doctorat de sociologie sous la direction de Bertrand Geay et soutenue à Amiens, Université de Picardie Jules Verne, le 11 mars 2013.

⁵³⁵ La loi n° 2005-380 du 23 avril 2005 d'orientation et de programme pour l'avenir de l'École.

⁵³⁶ Trois ministres ont successivement suivi la mise en place de la réforme du socle commun, de ses prémices jusqu'à l'adoption de la loi : Luc Ferry contribue à la mise en place du débat national, François Fillon en suit les débats au parlement, et Gilles de Robien en suivra l'application réglementaire.

⁵³⁷ La loi n°2013-595 du 8 juillet 2013 d'orientation et de programmation pour la refondation de l'école de la République.

par François Fillon à l'Assemblée nationale en procédure d'urgence. Le rapporteur de la commission des affaires culturelles saisie pour examiner le texte à l'Assemblée nationale est Frédéric Reiss, député UMP du Bas-Rhin, par ailleurs agrégé de mathématiques. L'Assemblée examine le texte en première lecture du 15 au 18 février 2005. Au Sénat, le texte est examiné du 15 au 19 mars 2005. Son rapporteur est Jean-Claude Carle, sénateur UMP de Haute-Savoie, agent technique de profession. Le texte définitif sera adopté à l'unanimité d'accord de la commission mixte paritaire réunie le 24 mars. Le conseil constitutionnel a été saisi et censurera deux dispositifs présentés à l'article 7 et à l'article 12. Le premier dispositif définissant les missions de l'école a été considéré comme ne relevant pas de la loi. Le second approuvait un rapport présenté en annexe qui n'avait pas été préalablement présenté au Conseil économique et social, et qui incitait à réduire le nombre d'options dans l'enseignement secondaire.

Les grands axes de la loi :

- La mise en place du socle commun de compétences et de connaissances.
- La mise en place de dispositifs visant à lutter contre l'échec scolaire et à favoriser l'accompagnement individuel des élèves, permettant d'atteindre l'objectif de réussite de tous d'une classe d'âge au niveau du baccalauréat : création du contrat individuel de réussite éducative, mise en place de soutien scolaire.
- Le renforcement de l'enseignement des langues par l'ajout d'une épreuve de langues vivantes au concours de recrutement des professeurs des écoles, le renforcement de l'enseignement de l'expression orale dans l'enseignement des langues au niveau secondaire.
- La réforme de la formation des enseignants, confiée à l'Université.
- La mise en place d'un effort de gestion pour le système éducatif, en mettant notamment en place des contrats d'objectifs entre les établissements scolaires et leur académie.
- Le Haut conseil de l'éducation est créé, visant à donner au gouvernement un avis sur les questions relatives à la pédagogie, aux programmes, à l'organisation du système éducatif à ses résultats.

La loi de 2013 pour la refondation de l'école de la République :

Elle est portée par Vincent Peillon, et promulguée le 8 juillet 2013. Une concertation préalable est également lancée au niveau national, avant que le projet de loi soit déposé à l'Assemblée nationale le 23 janvier 2013 en première lecture. Yves Durand, sénateur socialiste de Mayenne en est le rapporteur pour la commission des affaires culturelles de l'éducation saisie au fond sur le texte. Il est débattu en séance du 11 au 19 mars 2013. Au Sénat, le texte est examiné en première lecture du 21 au 24 mai 2013. Sa rapporteure est Françoise Cartron, sénatrice socialiste de Gironde. Une deuxième lecture dans chaque chambre aura lieu en juin 2013, et le texte sera adopté le 25 juin 2013. La loi est la première loi relative au système scolaire qui fera l'objet d'un suivi de son application de ses résultats par un comité dédié. Elle suscitera de vives oppositions concernant la mise en place de nouveaux rythmes scolaires impliquant l'ajout d'une demi-journée d'école.

Les grands axes de la loi :

- Le développement de la scolarisation des enfants de moins de trois ans.
- La prise en considération de la culture dans la définition du socle commun de connaissances et de compétences.
- La création d'un service public de l'enseignement numérique et de l'enseignement à distance.
- L'organisation de la scolarité sous forme de cycles.
- Le renforcement de l'enseignement des langues vivantes et l'obligation d'apprendre une langue vivante dès le CP.
- Le remplacement du cours d'instruction civique par un cours d'enseignement moral et civique.
- La création des ESPE, écoles supérieures du professorat et de l'éducation, pour renforcer et améliorer la formation des enseignants, aussi bien initiale que continue.

La démarche engagée dans ce chapitre est celle de « *penser la réforme* », selon l'expression employée par Philippe Bezes⁵³⁸, en rendant compte des stratégies de choix, par les initiateurs de réformes, de la revendication de l'appui sur des évaluations et les jugements des instances qui les produisent, selon l'expertise qu'elles permettent de conférer au processus de réforme. « *L'objectif des entrepreneurs de cause et de tous ceux concernés par un problème public est de lui voir apporter une solution. (...) C'est là l'opération de mise en politique publique* »⁵³⁹, nous dit Erik Neveu. Nous observerons, lors des deux moments réformateurs de 2005 et de 2013, et la façon dont les données produites pour l'action par l'OCDE avec PISA et par la DEPP sont utilisées à cette fin. Nous tenterons de comprendre dans quelle mesure cela contribue à les instituer leur rôle d'expert.

La mise en politique publique par le biais d'une réforme implique une série de séquences législatives, linéaires dans le temps. Il est tout d'abord nécessaire, pour le ministre qui porte une réforme, de la justifier. Cette justification est d'autant plus importante que l'histoire des réformes de l'éducation nationale est marquée par de nombreux échecs et renoncements, au point d'en faire aux yeux de certains auteurs « *une institution rétive à la réforme* »⁵⁴⁰. L'École se voit considérée comme « *bloquée* »⁵⁴¹, incapable d'être l'objet de réformes suffisamment structurantes compte tenu des oppositions, voire des affrontements, entre les acteurs du champ⁵⁴². L'étude des jeux de pouvoirs et d'acteurs, dans des travaux proches de la sociologie des organisations, donne à comprendre la difficulté à mener des réformes. Les jeux peuvent se jouer entre les acteurs de terrain, où il arrive que les intérêts des instituteurs et des professeurs⁵⁴³ ou ceux des syndicats d'enseignement entre eux⁵⁴⁴ s'opposent. Ils peuvent se jouer

⁵³⁸ Philippe Bezes, « Pour une sociologie des politiques de réforme de l'administration », dans *Réinventer l'Etat, les réformes de l'administration française (1962-2008)*, Paris, PUF, 2009, p.14.

⁵³⁹ Erik Neveu, « L'analyse des problèmes publics », *Idées économiques et sociales*, n°190, décembre 2007, p.15.

⁵⁴⁰ André Legrand, « L'éducation nationale, une institution rétive à la réforme ? », *Politiques et management public*, 2005.

⁵⁴¹ Pierre Clément, *thèse cit.*

⁵⁴² Voir par exemple Antoine Prost, *Histoire de l'enseignement en France, 1800-1967*, Paris, Armand Colin, 1968.; *Education, société et politiques. Une histoire de l'enseignement en France de 1945 à nos jours*, Paris, Le Seuil, 1992.; et Suzanne Citron, *L'école bloquée*, Paris, Bordas, 1971.

⁵⁴³ Jean-Marie Donegani et Marc Sadoun, « La réforme de l'enseignement secondaire en France depuis 1945 : analyse d'une non-décision », *Revue française de science politique*, vol.26, n°6, 1976.

⁵⁴⁴ Bernard Toulemonde, *Petite histoire d'un grand ministère*, Paris, Albin Michel, 1988.

également entre les syndicats et les grands corps de l'administration de l'Education nationale⁵⁴⁵, ou entre les partis politiques en lien avec les autres acteurs du champ, aux doctrines opposées sur l'objet des réformes envisagées. Face à ces oppositions régulières de fond, qui portent en elles la potentialité de mettre à l'arrêt la dynamique réformatrice souhaitée par les ministres, le pouvoir du chiffre devient un des outils favorisant le succès d'une réforme, ou au moins son adoption. Cela passe non seulement par le travail en amont des textes, mais aussi par l'explicitation lors des débats en séance du bienfondé de ce qui est proposé et de l'anticipation de la forme d'évaluation des résultats atteints, en ligne avec les principes de Nouveau management public régissant la modernisation de l'action publique. Le mécanisme de justification à l'œuvre dans les réformes auxquelles nous nous intéresserons ici a plusieurs types de destinataires, qui ne sont pas exclusifs les uns des autres. D'une part, il s'agit d'en expliquer le bienfondé aux citoyens et citoyennes, dans un souci de transparence de la décision publique et démocratique. Il s'agit ensuite de fournir à ceux qui en seront les publics-cibles – les membres de la communauté éducative pour les réformes qui nous intéressent ici – la preuve que cette réforme à venir aura un effet bénéfique sur le fonctionnement et les résultats du système sur lequel elle porte. La justification de la pertinence et de l'inévitabilité de la réforme s'adresse également aux parlementaires, en ce qu'elle permet de poser les pierres fondatrices du débat législatif sur le sujet en les matérialisant par un exposé des motifs de la loi.

Le second temps est donc celui du débat parlementaire en lui-même, la discussion autour des propositions de politiques publiques nouvelles ou réorientant les précédentes vers de nouveaux objectifs avec de nouveaux moyens. L'évaluation devient un outil mobilisé pour mesurer, sous la plume des rapporteurs des commissions saisies, la pertinence des propositions faites dans la réforme compte tenu du contexte dans laquelle elle est lancée, et appuyer les amendements considérés comme pertinents pour ajuster les moyens proposés par le ministre aux résultats attendus.

⁵⁴⁵ Frédéric Sawicki, « Pour une sociologie des problématisations politiques de l'école, *Politix*, n°98, 2012.

Le troisième temps d'une réforme dans lequel intervient l'expertise des évaluateurs des politiques publiques est celle de la construction de mécanismes d'évaluation a posteriori de l'application de la loi et des résultats qui y sont associés. Cette évaluation de l'atteinte des objectifs assignés à une réforme s'inscrit dans les pratiques de Nouveau Management Public liées au passage d'une obligation de moyens à une obligation de résultats. Dans le cas des réformes qui nous intéressent, les mécanismes d'évaluation des résultats des réformes de 2005 et de 2013 ont été intégrés parmi les sujets portés lors de l'examen des projets de lois au parlement.

Nous verrons donc dans une première partie comment les évaluations nationales de la DEPP et l'évaluation PISA de l'OCDE sont utilisées par les ministres pour construire l'argument fondateur de leur réforme, sans lequel le travail de discussion parlementaire autour des dispositifs proposés peut difficilement se mener jusqu'au bout. La sollicitation des évaluations n'est pas neutre⁵⁴⁶. D'une part, elle confère, selon les moments où elle a lieu et les interlocuteurs à qui elle est présentée, une certaine « valeur » à l'argumentaire. Nous regarderons les usages différenciés des évaluations de la DEPP et de PISA en fonction des moments de construction de la réforme, qui témoigne d'une vision spécifique du ministre des enjeux qu'il perçoit être ceux défendus par ses interlocuteurs. L'expert n'est pas alors légitime en soi, mais acquiert cette qualité indispensable à la valeur de l'expertise qu'il produit en fonction de celui qui la reçoit. D'autre part, véritables « traceurs du changement ⁵⁴⁷ », les évaluations PISA mettent en exergue, par l'intensité avec laquelle elles sont sollicitées par les ministres réformateurs et les contributeurs à la construction de la réforme, la dynamique d'euphémisation puis d'internationalisation des enjeux de l'éducation. Cela placera par construction la DEPP en dehors de la liste des instances expertes sur les problèmes publics en éducation compte tenu du cadre dans lesquels sont travaillés ces enjeux.

Nous analyserons dans une deuxième partie les usages des évaluations nationales et de PISA dans les deux grands débats parlementaires lors des réformes de

⁵⁴⁶ Nous nous inscrivons ici dans la suite des travaux Pierre Lascoumes et Le Galès, *Gouverner par les instruments*, Paris, Presses de Sciences Po, 2005.

⁵⁴⁷ Bruno Palier, « Chapitre 7 : Les instruments, traceurs du changement. La politique des retraites en France », Pierre Lascoumes éd., *Gouverner par les instruments*. Presses de Sciences Po, 2005.

2005 et de 2013. Nous chercherons à comprendre comment l'usage des évaluations contribue à l'attribution d'un rôle d'expert en évaluation de la politique éducative. Ce rôle d'expert ne consiste pas seulement à « produire » sur commande des savoirs mobilisés pour l'action, selon la formule employée par Corinne Delmas dans son étude du concept d'expertise en action publique⁵⁴⁸. Il a aussi pour effet de révéler des savoirs préexistants dans le débat public, parfois déjà identifiés comme problème public au sens où ils avaient fait l'objet de tentatives de réorientation antérieure de la politique éducative. Nous verrons que les opposants politiques des théories économiques et de la vision du développement traditionnellement portés par l'OCDE, principalement à gauche de l'échiquier politique, vont finalement s'appuyer sur l'outil pour valider l'acceptabilité du constat de l'aggravation des inégalités scolaires en France au contraire de ses partenaires internationaux. Ils fondent ainsi leur réforme et permettent de justifier la pertinence des solutions qu'ils souhaitent y apporter. L'utilisation aux différents moments de la réforme des résultats de PISA a construit, au fur et à mesure de la période, de nouveaux destinataires à convaincre du bienfondé aussi bien des objectifs assignés que des moyens mis en œuvre pour les atteindre : les pays partenaires de la France et l'OCDE en elle-même. L'OCDE devient l'évaluatrice *per se* de la réforme pour la refondation de l'école, réduisant dans le même temps le pouvoir d'influence de la DEPP dans le pilotage de l'action publique en éducation. L'enjeu est de montrer, par les évaluations PISA et à l'OCDE, que la France a fait les bons choix de réforme et a réussi à les mettre en œuvre. La DEPP se voit alors attribuer le rôle de passeur⁵⁴⁹ des données et recommandations issues de PISA auprès des directions opérationnelles du ministère et auprès des enseignants et de leurs formateurs. Forte des liens qu'elle a tissés avec le corps enseignant, notamment grâce aux trajectoires professionnelles des personnes nommées à sa tête à partir des années 2010, son expertise devient davantage une maîtrise technique des chiffres produits par d'autres pour l'action publique en éducation.

⁵⁴⁸ Corinne Delmas, *Sociologie politique de l'expertise*, La Découverte, 2011.

⁵⁴⁹ Sur la figure du « passeur » contribuant, par ses attributions professionnelles ou par ses convictions, à la circulation de normes et de politiques publiques, voir par exemple Fabien Jobard, et al. « Sociologie politique des passeurs. Acteurs dans la circulation des savoirs, des normes et des politiques publiques », *Revue française de science politique*, vol. 70, n° 5, 2020.

1) La construction de l'argument fondateur de la réforme : quels évaluateurs pour quelle expertise ?

Loin de ne proposer qu'une amélioration linéaire du système éducatif français, la réforme telle que l'entendent les ministres qui portent celles de 2005 et de 2013 déploie une vision holistique des problèmes d'action publique en éducation auxquels apporter une réponse, et se tourne à la fois vers le passé et vers le futur. Vers le passé, car elle vise à retrouver l'essence de l'école républicaine et à transformer le système éducatif en profondeur pour lui permettre de retrouver ses missions initiales – selon ce que chaque ministre met derrière ce terme. Vers le futur, car elle s'inscrit dans une démarche de modernisation, que cela soit entendu au sens « managérial » du terme⁵⁵⁰, c'est-à-dire comme la recherche d'un usage optimal des ressources pour atteindre la meilleure performance, en favorisant la responsabilité individuelle et l'évaluation des résultats, ou au sens « politico-administratif » du terme, c'est-à-dire la réorientation nécessaire de l'action publique pour lui permettre de conserver sa pertinence dans un contexte futur indéterminé, quelle que soit finalement sa forme.

Les conditions de réussite d'une réforme sont multiples. Relevant pour certains du contexte social et économique dans lequel elles sont lancées⁵⁵¹, de l'adéquation avec un moment propice à la réception des idées qu'elles portent par les citoyens et les médias, de la personnalité des ministres qui les mènent⁵⁵², elles sont aussi dépendantes de la capacité à développer un discours favorisant l'adhésion de ceux qui en seront les cibles, démontrant son bienfondé et l'ampleur et l'ambition des solutions proposées aux problèmes publics traités. Nous nous concentrerons dans cette première partie sur la constitution des cadres argumentatifs portant sur la nécessité de la réforme, qui en facilitent le lancement et en favorisent l'aboutissement. Nous observerons comment plusieurs des résultats de PISA ont d'abord été intégrés, avec les évaluations de la

⁵⁵⁰ Philippe Bezes, *Réinventer l'Etat, les réformes de l'administration française (1962-2008)*, Presses Universitaires de France, 2009.

⁵⁵¹ John Kingdon, *Agendas, Alternatives and Public Policies*, New-York, Harper Collins, 1984.

⁵⁵² Ce fût le cas notamment avec l'arrivée de Claude Allègre du ministère. Ministre aux déclarations suscitant l'émoi du corps enseignant sur l'incapacité de l'Education nationale à se moderniser et à rationaliser ses coûts de fonctionnement, il finira par démissionner à la demande de Lionel Jospin en mars 2000, à la suite d'un énième conflit avec les enseignants ayant généré des manifestations d'ampleur.

DEPP, dans un panel d'arguments multiples utilisés par un gouvernement politiquement à droite, pour recentrer l'évaluation des résultats du système éducatif autour de l'acquisition d'un socle commun de compétences. Il devient ensuite, par sa capacité à être largement entendu par tous les partis quel que soit la couleur politique de celui qui l'utilise, l'argument principal d'un gouvernement de gauche habituellement moins adepte des constats et des préconisations de l'OCDE, pour justifier la concentration des efforts réformateurs autour de la réduction des inégalités scolaires.

a) Inscrire dans le marbre le concept du socle commun, inspiré de la vision portée par l'OCDE des objectifs de l'éducation

Le 7 décembre 2004, sur France 2 lors du journal du matin, le ministre de l'Éducation nationale, François Fillon, est interrogé sur les résultats de l'enquête PISA tout juste publiés. Il constate la stagnation voire « l'érosion » des résultats scolaires des jeunes Français. « *C'est la raison pour laquelle, je propose une loi pour réformer l'école* », annonce-t-il. « *Depuis plusieurs années, on sent que nos performances plafonnent, voire même qu'elles sont en lente érosion par rapport aux autres pays européens, et cela malgré les efforts considérables qui ont été faits depuis 15 ans, notamment en termes de moyens (...). Cela signifie bien qu'il faut aller maintenant au-delà de ces questions de moyens et modifier notre façon d'enseigner, notre organisation scolaire pour avoir de meilleurs résultats* ». La réforme qu'il propose est matérialisée par le projet de loi d'orientation et de programmation sur l'avenir de l'école, que nous appellerons par économie ici la loi de 2005. La question des moyens alloués au système éducatif est balayée par le ministre d'emblée, celui-ci préférant insister dans son moment de communication sur le cœur de la réforme qu'il souhaite porter : le changement de façon de concevoir le résultat attendu de l'éducation, et la constitution d'un système d'évaluation permettant de regarder comment ce résultat est effectivement atteint. Le socle commun de compétences comme objectif minimum à atteindre est mis au centre de la réforme, et sa pertinence est justifiée à la fois par le constat, alimenté par la DEPP, d'un nombre important de décrocheurs et de jeunes sortis sans qualification du système scolaire, et par l'exigence

internationale de raisonner en termes d'acquisition de compétences clés qui permettront aux élèves devenus actifs sur le marché du travail de trouver avec succès un emploi.

i. Légitimité des uns donc expertise du groupe: la reconnaissance des travaux de la DEPP à travers la commission Thélot

La réforme est préparée dès l'été 2003 par l'annonce le 3 juillet, devant les parlementaires, de la mise en place d'un mode nouveau de concertation, un grand débat national sur l'avenir de l'école, auquel sont invités à s'exprimer tous les Français⁵⁵³. Ce choix, revendiqué par Luc Ferry et Xavier Darcos dans leur intervention devant la représentation nationale, est fait pour rapprocher la réforme engagée des souhaits exprimés par la communauté éducative au sens large, et ainsi désamorcer les procès en déconnexion qui pourraient être faits par les représentants syndicaux des corps d'enseignement et les acteurs de terrain⁵⁵⁴. Par ailleurs, les prémices de la réforme sont lancées alors que de nombreux conflits sociaux ont jalonné l'année scolaire 2002-2003, liés aux mesures de rigueur budgétaire prises pour le budget de l'Education nationale et au projet de réforme des retraites⁵⁵⁵ contre lequel les syndicats de l'enseignement se sont fortement mobilisés. En rendant la préparation de la réforme participative, et malgré le cadrage préalable important des thèmes de travail dont la concertation fait l'objet de la part du ministre, la création d'une commission nationale du débat sur l'avenir de l'école permet de déplacer les discussions en dehors du champ des conflits sociaux de la période et, selon les mots

⁵⁵³ Le débat national sur l'avenir de l'école, qui a lieu de novembre 2003 à janvier 2004, voit organisées en un temps record 26 000 réunions, faisant participer plus d'un million de personnes. Les synthèses des réunions ont été mises en ligne, et le site sur lequel elles ont été publiées a été visité 400 000 fois. 15 000 internautes ont déposé quelques 50 000 messages. Des lettres de particuliers et des contributions de collectivités et d'acteurs associatifs ont part ailleurs été reçues en nombre. Les plus présents, selon le rapport rendu par la commission Thélot, intitulé *Le miroir du débat*, ont été les enseignants, puis « à un degré moindre celle des parents, beaucoup moins celle des autres acteurs, notamment des jeunes ». (p.13).

⁵⁵⁴ Pierre Clément, *op. cit.*

⁵⁵⁵ Bertrand Geay, « Le « tous ensemble » des enseignants », dans Sophie Bérout et René Mouriaux, *L'année sociale*, Syllepse, Paris, 2004.; et Bertrand Geay, *Le syndicalisme enseignant*, Paris, La Découverte, 2005.

des ministres Ferry et Darcos, d'élaborer « *un diagnostic objectif et partagé sur la situation de l'école* ⁵⁵⁶ ». Le dialogue est noué pour identifier les problèmes structurels du système éducatif français, à la lumière des apports de chaque participant selon la singularité de son regard et de ses expériences. Charge ensuite au gouvernement de proposer les solutions adéquates aux constats posés, une fois le rapport de concertation remis par la commission ad hoc. L'objectivité des conclusions rendues est garantie par la nomination des responsables du travail menée par la commission et notamment de son président, Claude Thélot.

Claude Thélot est nommé par les ministres Luc Ferry et Xavier Darcos pour ses compétences reconnues par tous en matière d'évaluation et de diagnostic sur l'état du système éducatif que son parcours avait confortées. Polytechnicien de formation, passé par l'ENSAE avant de commencer sa carrière à l'INSEE, directeur de la DEPP pendant sept ans de 1990 à 1997, conseiller-maître à la Cour des comptes par la suite puis Président du Haut Conseil de l'évaluation de l'école, il est un fin connaisseur des rouages de la haute-administration. Proche du monde universitaire, il est également l'auteur de plusieurs ouvrages sur l'histoire des politiques éducatives françaises et leur évaluation⁵⁵⁷ aussi bien du point de vue de leurs résultats en termes de reproduction sociale ou d'accès au marché du travail, que du point-de-vue de leur méthode et de l'« *étroite interaction entre les questions que l'on se pose, et donc le type de réponse que l'on apporte, et les instruments d'observation et d'analyse qui sont retenus* »⁵⁵⁸. Même s'il fût proche de Lionel Jospin lorsqu'il devint ministre de l'Éducation nationale, notamment par l'adéquation de vision de l'apport de l'évaluation et de l'outil statistique pour le pilotage de l'action publique, ses différents postes en dehors de tout cabinet ministériel en font par ailleurs un expert transpartisan. Cette connaissance pluridisciplinaire de l'administration et du système éducatif français, à la fois théorique et opérationnelle

⁵⁵⁶ Compte rendu de l'audition publique de Luc Ferry et Xavier Darcos devant la commission des affaires culturelles du Sénat, 3 juillet 2003.

⁵⁵⁷ Voir notamment Claude Thélot, *Tel père, tel fils ?*, Paris, Dunod, Bordas, 1982 ; Claude Thélot, *L'évaluation du système éducatif*, Paris, Nathan, 1994 ; Claude Thélot et Olivier Marchand, *Le travail en France, 1800-2000*, Nathan, collection Essais et Recherches, Paris 1997 ; Claude Thélot et Philippe Joutard, *Réussir l'école*, Paris, Ed. du seuil, 1999 ; Claude Thélot, *Les écrivains racontent l'école*, Delagrave, 2001 ; ou encore Claude Thélot, *L'origine des génies*, Paris, Le Seuil, 2003.

⁵⁵⁸ Claude Thélot, *Tel père, tel fils ?*, Paris, Dunod, Bordas, 1982.

du fait de son passage par la haute-administration, rend sa nomination à la tête de la commission du débat national sur l'avenir de l'école tout à fait consensuelle au sein des participants, qu'ils soient représentants syndicaux ou parlementaires, enseignants ou chercheurs⁵⁵⁹.

Cette nomination est vécue par la DEPP comme une reconnaissance de la qualité du travail mené par la direction du ministère depuis une décennie. L'un de nos interlocuteurs nous confiait en entretien le retentissement positif qu'avait pu avoir la présidence de Claude Thélot sur la légitimité des travaux de la DEPP, en interne et auprès des publics cibles des publications de la DEPP.

« C'est Claude Thélot qui a quand même été nommé président de la commission. Alors vu ce qu'il a apporté à la DEPP et sa position sur le rôle de la DEPP auprès du ministre, c'est quand même un signe fort. Et ça a permis aussi de faire forcément connaître un peu son parcours, et du coup aussi notre travail. »⁵⁶⁰

La nomination de Claude Thélot à la tête de la commission qui portera son nom vient apporter reconnaissance et légitimité aux administrations dont il avait la charge quelques années auparavant, dont la DEPP. Elle vient renforcer symboliquement la validité des travaux de la DEPP auprès des ministres qui mèneront le débat national et l'examen de la loi de 2005, après plusieurs années de relations compliquées sous le ministère Allègre qui se sont d'ailleurs achevées par la démission de Thélot. Elle permet aussi de les populariser potentiellement auprès de l'ensemble des participants peu connaisseurs des travaux d'évaluation du système éducatif menés par la DEPP. En confiant à Claude Thélot les rennes de la commission du débat national sans opposition notable, le gouvernement acte et les participants reconnaissent non seulement l'impartialité de Thélot et son indépendance du jeu politique, mais aussi la qualité des compétences acquises et mises en pratique pendant les années passées à la DEPP.

Cette légitimité scientifique et administrative est reprise à son compte par les équipes de la DEPP, dont la valeur des travaux pour l'orientation de la politique

⁵⁵⁹ Pierre Clément, *thèse. cit.*

⁵⁶⁰ Entretien avec Jean-Pierre Jeantheau, Chargé d'étude, National Project Manager, DEPP, 21 mars 2019.

éducative est renforcée par leur intégration dans le processus de construction de la réforme à travers leur usage lors de la préparation des débats sur l'avenir de l'école. On comprend bien, au regard de l'ampleur des chiffres qui décrivent le déroulement du débat national et ses dizaines de milliers de contributions orales et écrites, que l'organisation du cadre du débat et notamment des questions posées et des éléments de contexte fournis aux participants en libre accès est fondamentale pour rassembler les propositions, canalisées dans des directions prédéfinies. Les travaux mobilisés par Claude Thélot pour préparer le grand débat national sur l'avenir de l'école sont, pour ce qui concerne les données statistiques, largement issues des évaluations menées par son ancienne direction du ministère de l'Éducation nationale. En octobre 2003, le Haut Conseil de l'évaluation de l'école, dont Thélot a été président de 2000 à 2003, publie un document général préparatoire au débat national sur l'avenir de l'école, intitulé *Eléments pour un diagnostic sur l'École*. Il est rédigé par Jean-Claude Hardouin, professeur, économiste, et ancien conseiller de François Bayrou, André Hussenet et George Septours, inspecteurs généraux de l'éducation nationale, et fait intervenir Norberto Bottani, à l'époque directeur du service de la recherche en éducation à Genève après avoir travaillé plus de 20 ans au sein du CERI de l'OCDE, pour apporter un regard international aux constats dressés. L'objectif de ce rapport de plus de 150 pages n'est pas de donner à voir une photographie exhaustive du fonctionnement du système éducatif et des débats actuels en matière de politique éducative, mais bien de fournir un bagage commun et synthétique sur l'école des années 2000 aux participants qui en exprimeraient le besoin⁵⁶¹. Pour cela, le HCÉE s'appuie sur ses propres travaux et sur les travaux menés par ailleurs par les auteurs du rapport, mais également sur toute une série d'évaluations et d'indicateurs nationaux et internationaux dont il donne les producteurs :

« Nous avons utilisé nos propres observations, et nous nous sommes appuyés autant que la matière le permettait, sur les données chiffrées des publications officielles de l'OCDE, des services de l'Union Européenne, de l'INSEE, de la direction de l'évaluation et de la

⁵⁶¹ HCÉE, *Eléments pour un diagnostic sur l'École. Document général préparatoire au débat national sur l'avenir de l'école*, 2003.

prospective du ministère en charge de l'éducation. Nous avons réputé indiscutables ces données. »⁵⁶²

Même si la DEPP est citée en dernier, ses données et ses évaluations sont utilisées tout au long des chapitres du rapport. Elle apparaît plus d'une dizaine de fois dans le texte, et est convoquée pour donner à voir certains traits caractéristiques du système éducatif français y compris de façon longitudinale (la part de l'enseignement privé par exemple, depuis les années 1980, ou la part de jeunes commençant un second cycle professionnel après la classe de 3^e), ainsi que des évaluations des résultats de mesures mises en place lors de précédentes réformes et répondant à des problèmes publics présentés comme des enjeux de société (l'effet de la scolarisation à deux ans sur l'acquisition des savoirs scolaires en primaire, ou encore les trajectoires scolaires des enfants scolarisés en Zone d'Education Prioritaire).

PISA apparaît quant à lui déjà malgré sa mise en place plus récente à cinq reprises dans le rapport, principalement pour donner des informations sur des résultats de l'évaluation de dimensions extra-disciplinaires comme le climat au sein de la classe et le sentiment de soutien perçu par les élèves de la part de leurs professeurs. La question de l'effet du milieu socio-économique des élèves sur leurs performances au test, et la comparaison des résultats des élèves français par rapport aux autres pays, qui deviendra l'élément central de la réforme de 2013 lancée par Vincent Peillon et sur laquelle nous reviendrons, est déjà mentionnée⁵⁶³ sans que cela ne donne lieu à une reprise particulière de sujet lors des ateliers du débat national. Bien que les mentions des travaux de PISA soient limitées dans les documents préparatoires au débat national, plusieurs indices témoignent du statut d'expert qu'occupe déjà l'OCDE lorsqu'il s'agit de débattre des résultats des systèmes éducatifs. D'abord, ses données sont considérées en postulat de départ « indiscutables ». Deux ans seulement après les publications des premiers résultats dont la méthode avait suscité des critiques y compris au sein de la DEPP et des acteurs du champ éducatif, la validité des données de PISA semble acquise et leur usage tout à fait approprié à la réflexion de fond sur le

⁵⁶² HCÉE, *op. cit.* p.3.

⁵⁶³ HCÉE, *op.cit.* p.31.

bienfondé d'une réforme systémique. Si l'on regarde plus largement la présence de l'OCDE dans le rapport du HCÉE, on voit que les données PISA et les données descriptives qu'elle produit par ailleurs annuellement sur le système éducatif français dans *Regards sur l'éducation* sont citées plus d'une vingtaine de fois. Elles font même l'objet de comparatifs sous forme de graphiques mettant en regard les données du système français par rapport à la moyenne des pays de l'OCDE, sur les temps d'enseignement par discipline, la part de la population ayant terminé le lycée, ou le taux d'obtention par génération d'un diplôme d'enseignement supérieur.

Ce rôle de photographe du système éducatif attribué à l'OCDE se double d'un rôle d'expert dans la construction des orientations qui figureront dans le rapport de la commission Thélot remis au ministre Fillon pour contribuer à rédiger son projet de loi. L'OCDE est citée parmi les organisations, experts et personnalités entendus⁵⁶⁴, contrairement à la DEPP qui n'apparaît pas. Au sein de la commission se trouvent par ailleurs de façon permanente deux chercheurs étrangers, caractérisant l'ouverture voulue par Claude Thélot, qui souhaitait déplacer le cadre du débat national en dehors des espaces d'échanges habituels afin de s'éloigner tant que possible des discours convenus et des potentielles querelles d'initiés. Parmi eux, Tom Schuller, professeur doyen de la faculté de formation continue à l'université Birkbeck, est également le responsable du CERI de l'OCDE. Du côté de la DEPP, si ce n'est la présidence de Claude Thélot dont l'attribution tient en grande partie à la finesse de son expertise exercée à la tête de la direction pendant sept années, aucun autre de ses agents n'est membre de la commission du débat national sur l'avenir de l'école, et cela bien que la démarche d'évaluation continue des résultats de la réforme soit l'un des axes clés exposés par François Fillon. La direction retrouve par cette absence son statut de direction administrative du ministère, agissant sur commande du ministre pour mettre en place a posteriori les évaluations pré-pensées pour correspondre aux besoins politiques de retour sur l'application des réformes. Experte au sens où elle a produit par ses missions quotidiennes une partie des savoirs mobilisés pour l'action réformatrice⁵⁶⁵, elle n'est paradoxalement pas reconnue comme maîtresse de ses

⁵⁶⁴ Annexe IV, p.150.

⁵⁶⁵ On retrouve ici le concept tel qu'entendu par Corinne Delmas.

travaux ni intéressante à solliciter pour les porter à la connaissance des débatteurs, alors même que Claude Thélot, qui a façonné la DEPP telle qu'elle a existé dans les années 1990, a justement été nommé pour sa capacité à faire le lien entre évaluation, analyse et construction de la politique éducative⁵⁶⁶.

ii. PISA et le socle commun de compétences : entre lien explicite et omission stratégique

Si les rapports publiés par la commission Thélot à l'issue du grand débat national ne feront pas l'objet en soi d'un usage extensif par le ministre Fillon lors de la préparation de l'examen parlementaire de son texte, les sujets qu'ils abordent auront été portés pendant plusieurs mois auprès de l'ensemble de la communauté éducative. Le concept de socle commun de compétences, jusqu'ici discret dans les débats entre les acteurs du système éducatif, fait son entrée dans le champ des possibles des orientations à venir d'un système éducatif français jusqu'alors particulièrement attaché aux contenus de programmes définis précisément au niveau national. Stratégiquement, la conception les grandes questions posées lors des ateliers du débat national a intégré la mise en place du socle commun comme un prérequis à la réforme, sur laquelle les participants devaient se prononcer et proposer des nouvelles orientations. La question « faut-il un socle commun ? », induisant déjà une réponse binaire laissant le champ ouvert à l'adhésion au principe du socle, a été remplacée par « quel socle commun faut-il ? », partant du constat de sa mise en œuvre à venir et interrogeant sur ses formes souhaitables, et par dérivation les modalités de son évaluation. Pour Claude Thélot lui-même, interrogé sur l'origine des débats sur le socle par Pierre Clément, le bienfondé de la recherche d'un socle commun n'est pas porté au débat car le concept avait été en amont largement accepté par la majorité des membres de la commission.

⁵⁶⁶ Lors de l'entretien qu'il avait accordé à Pierre Clément lors de ses travaux de recherche doctorale, Claude Thélot, interrogé sur le moment de sa nomination par Lionel Jospin à la tête de la DEP en 1990 avait déclaré : « C'est un poste admirable parce qu'on est vraiment à cheval entre l'analyse et la politique. Comme directeur d'administration centrale vous êtes à la fois conseiller du ministre et responsable de l'ensemble de la politique éducative qui est conduite », p.580.

« (Il lit) "Quel socle commun de connaissances ?" ... C'est-à-dire, qu'on a voulu que le débat ne se concentre pas sur "faut-il un socle ou pas ?" ... Parce que, à mon avis, on ne s'en sort pas avec des questions posées de cette manière. Comme pour l'ensemble des questions qu'on a voulu poser d'ailleurs, les 22, on a plutôt essayé de privilégier des questions un peu opérationnelles. Ça c'était vraiment une exigence que j'avais marquée. C'est-à-dire essayer de dire ce qu'il faut mettre dans le socle, plutôt que de dire : faut-il ou non raisonner en termes de socle ? »⁵⁶⁷

Le socle commun, objet phare de la réforme dès la préparation du débat national devient également le point central de la communication ministérielle à partir du lancement du processus législatif. Le communiqué de presse du Conseil des ministres du 12 janvier 2005, qui annonce la présentation de la réforme au gouvernement, présente les trois principales actions conçues pour permettre d'améliorer le niveau de formation des élèves français. La loi doit avant tout permettre de fournir à tous les élèves un socle de connaissances et de compétences jugées indispensables à la sortie du système scolaire. Elle pose également le principe de renforcer l'accès à l'enseignement supérieur avec l'objectif de mener la moitié d'une classe d'âge à ce niveau d'enseignement. Elle place enfin la modernisation de l'Education nationale en principe directeur de l'organisation du système éducatif. Les deux derniers axes du texte de loi sont dès le départ éclipsés par le concept nouveau de socle commun. Présenté dans l'article 9 du texte de loi, il prend dès les premiers moments de communication du ministre l'ascendant sur les 88 autres articles⁵⁶⁸. Le socle commun concentre les efforts d'explicitation et de justification de la réforme et plusieurs lignes argumentatives sont construites pour souligner la nécessité de l'introduire comme objectif premier de l'Education nationale. Le socle se voit ainsi défini par le ministre Gilles de Robien, qui prend la suite de François Fillon, comme « *un acte fondateur pour notre école, un moment exceptionnel dans l'histoire scolaire, sans équivalent depuis les lois de Jules Ferry qui ont instauré l'instruction gratuite, laïque et obligatoire et en ont précisé les contenus* »⁵⁶⁹. Au-delà de la glorification politique de la réforme mise en œuvre par les

⁵⁶⁷ Entretien avec Claude Thélot mené par Pierre Clément dans le cadre de sa thèse, cité p.588.

⁵⁶⁸ Pierre Clément, *op. cit.*

⁵⁶⁹ *Ecole et collège : tout ce que nos enfants doivent savoir*, CNDP/XO éditions, 2006, p.7-8.

ministres successifs du gouvernement Raffarin, la prise de parole de Gilles de Robien exprime l'importance accordée à l'inscription dans la loi du concept de socle commun, comparé aux principes des lois fondatrices de l'école de la République plus d'un siècle plus tôt. Avant même de passer par l'instrument législatif, qui lui confèrera la légitimité qu'apporte la validation par la représentation nationale de son caractère d'intérêt public⁵⁷⁰, le concept de socle commun doit être légitimé rationnellement par un arsenal argumentatif irréfutable. Pour cela, la stratégie de communication du ministre qui lancera le débat parlementaire sur la loi de 2005, François Fillon, se dote d'un ensemble d'indicateurs de résultats combinés à des propositions de réorientation de la politique éducative. La communication ministérielle autour de la légitimité du concept de socle commun et de la nécessité de le replacer au centre du système éducatif adapte les arguments issus des différentes évaluations nationales et de PISA à l'auditoire, sans qu'aucune ne prenne encore le pas sur les autres.

La force de l'argument PISA réside encore une fois dans sa dimension intrinsèque internationale, qui place l'évaluation dans un cadre spatial adapté au travail de la France en Europe pour construire un référentiel commun des enjeux à venir en matière d'emploi. La conceptualisation du socle commun n'est pas étrangère, dans la bouche même du ministre Fillon, aux recommandations internationales formulées par plusieurs institutions européennes de travailler sur ce qu'elles appellent les compétences clés, qui seraient elles-mêmes dérivées des évaluations menées par l'OCDE et PISA en particulier⁵⁷¹. Le processus de construction du principe de socle commun est ainsi a posteriori détaillé par le cabinet de François Fillon, en réponse à

⁵⁷⁰ Voir la typologie des instruments d'action publique formalisée par Charlotte Halpern et Patrick Le Galès, « Pas d'action publique autonome sans instruments propres. Analyse comparée et longitudinale des politiques environnementales et urbaines de l'Union européenne », *Revue Française de Science Politique*, 61 (1), 2011.

⁵⁷¹ Au-delà du lien fait par les décideurs politiques dans leurs discours de justification de leurs projets de loi, le sujet a été abordé par la recherche pour identifier l'influence possible de l'OCDE sur la construction de concepts clés repris au niveau européen, comme la littératie ou les compétences clés. Sur ce sujet voir par exemple Véronique Leclercq, « La notion de littéracie dans le champ des formations de base pour adultes : référence explicite ou influence diffuse ? », *Spirale, revue de recherches en éducation*, n°53, 2014 ; Jean-Claude Coulet, « Les notions de compétences clés : l'éclairage d'un modèle théorique fondé sur l'analyse de l'activité », *Activités*, 13-1, 2016.

une question du député UMP Denis Merville en mai 2006, inquiet de voir la définition du socle être en inadéquation avec la définition européenne des compétences clés.⁵⁷²

« La définition du socle commun de connaissances et de compétences prend appui sur la proposition de recommandation du Parlement européen et du Conseil de l'Union européenne en matière de « compétences clés pour l'éducation et l'apprentissage tout au long de la vie ». Elle se réfère aux évaluations internationales, notamment au programme PISA (Programme for International Student Assessment) qui propose une mesure comparée des connaissances et des compétences nécessaires tout au long de la vie. »

Sans toutefois correspondre mots pour mots aux principes de la méthode d'évaluation de PISA puisque la question de l'évaluation des élèves par rapport à un programme n'est pas évacuée – les compétences sont jointes à un socle de connaissances -, la réforme proposée par François Fillon adopte une logique similaire de transversalité disciplinaire des compétences à maîtriser à la fin de la scolarité. Elle est inspirée, de ses propres mots, par les travaux européens sur les compétences clés, reprenant dans le processus de réflexion les travaux engagés dans le cadre du lancement de PISA et par le programme DeSeCo (Définition et Sélection des Compétences clés) de l'OCDE cherchant à définir « *le cadre conceptuel pour orienter le développement à long terme des évaluations et l'extension à des nouveaux domaines de compétences* »⁵⁷³. Cette origine internationale du socle transparaîtra assez clairement lors de la publication du détail de ce qui constitue les compétences attendues des élèves dans le socle commun. Définies dans un décret paru le 12 juillet 2006, elles sont presque identiques dans leur formulation à celles du cadre de référence européen travaillé depuis la fin des années 1990 à la lumière des travaux de l'OCDE. Les compétences clés sont au nombre de sept dans la loi française contre huit au niveau européen. Cette légère différence ne provient pas d'un désaccord français sur l'importance d'une des compétences clés européennes. Elle résulte tout simplement du regroupement de deux d'entre elles.

⁵⁷² Question écrite n°93873 posée par Denis Merville, UMP, publiée au journal officiel de l'Assemblée Nationale le 9 mai 2006.

⁵⁷³ OCDE, *Définition et Sélection des Compétences clés*, 2005, p.3.

Tableau 3. Comparaison des compétences clés entre le document cadre du référentiel européen et le décret relatif à la loi de 2005

Compétences du socle de la loi de 2005⁵⁷⁴	Compétences clé du référentiel européen⁵⁷⁵
maîtrise de la langue française	communication dans la langue maternelle
pratique d'une langue vivante étrangère	communication dans une langue étrangère
principaux éléments de mathématiques et culture scientifique et technologique	culture mathématique et compétences de base en sciences et technologies
maîtrise des techniques usuelles de l'informatique et de la communication	culture numérique
culture humaniste	esprit d'entreprise sensibilité culturelle
compétences sociales et civiques	compétences interprofessionnelles, interculturelle compétences sociales et civiques
autonomie et initiative des élèves	apprendre à apprendre

Il est intéressant de noter ici que les compétences du socle souhaitées par François Fillon sont déjà listées dans leur majorité avant même l'examen de la réforme par l'Assemblée nationale et le Sénat. Le point central de la réforme est largement formalisé avant même d'être soumis au travail de la représentation nationale, et demeurera presque inchangé après son passage par la navette parlementaire, au moment de l'adoption définitive de la loi⁵⁷⁶.

Les références à l'Europe dans les espaces de communication de François Fillon sont régulières pour justifier la mise en œuvre du socle commun de compétences. Certains concepts clés comme certains des objectifs présentés par le ministre dans ses 14 propositions pour la refondation de l'école sont en effet directement issus des discussions européennes, et notamment des objectifs actés dans le cadre de la stratégie de Lisbonne. Il affirme ainsi vouloir mener la moitié d'une classe d'âge vers un diplôme de l'enseignement supérieur à l'horizon 2015, répondant à l'invitation du

⁵⁷⁴ Ministère de l'Éducation nationale, loi d'orientation sur l'École. 14 propositions de François Fillon, novembre 2004, p. 7.

⁵⁷⁵ Parlement européen, Recommandation du Conseil européen et du Parlement européen, 2006.

⁵⁷⁶ Pierre Clément, *thèse. cit.*

Conseil européen de réduire de moitié le nombre des 18-24 ans ne poursuivant pas leurs études ou leurs formations après le cycle secondaire. En reliant ces indicateurs de résultats chiffrés avec l'objectif figurant dans l'exposé des motifs de la proposition de loi, de « *répondre aux besoins de la société et de l'économie dans les prochaines décennies* » en « *favorisant la réussite de tous les élèves* »⁵⁷⁷, le ministre acte la diffusion dans une loi nationale d'un standard associé aux travaux de l'OCDE : celui de la définition d'un socle commun de compétences pour aboutir à un horizon commun sur le champ de l'emploi⁵⁷⁸.

Le lien souvent revendiqué par François Fillon entre PISA et sa conception du socle commun de compétences, une fois la réforme adoptée et les décrets parus, n'a cependant pas toujours été mis en exergue de façon systématique avant l'examen en séance du projet de loi. Il a davantage été employé en fonction du public-cible prévu, afin de maximiser sa force argumentative, et remplacé par d'autres évaluations quand cela a été jugé nécessaire. Cela est d'autant plus frappant que le texte est présenté en Conseil des ministres et annoncé à la presse tout juste un mois après la publication des résultats de PISA 2003. L'argument PISA, employé par le ministre en décembre 2004 sur France 2⁵⁷⁹ pour inscrire sa réforme dans l'actualité brûlante des résultats moyens des élèves français au test, n'a pas toujours été jugé suffisamment légitime pour convaincre l'ensemble de son auditoire. L'idée de transférer des concepts internationaux par l'intermédiaire d'une réforme de fond du système éducatif français n'a en effet pas nécessairement été systématiquement l'angle d'approche le plus efficace pour convaincre, témoignant d'une légitimité encore en construction de l'OCDE comme experte dans la construction des orientations de la politique éducative nationale, y compris pour un gouvernement de droite proche des théories portées par l'organisation internationale de façon générale.

Au moment d'intervenir dans l'émission de grande écoute *100 minutes pour convaincre* le 18 novembre 2004, les références à l'Europe se font notamment moins

⁵⁷⁷ Exposé des motifs du projet de loi pour l'avenir de l'école, présenté au nom de M. Jean-Pierre Raffarin, Premier ministre, par M. François Fillon, Ministre de l'éducation nationale, de l'enseignement supérieur et de la recherche, le 12 janvier 2005.

⁵⁷⁸ Institut national de recherche pédagogique, *Standards, compétences de base et socle commun*, décembre 2005, rédacteur Olivier Meunier.

⁵⁷⁹ Voir supra.

nombreuses. Le discours du ministre est recentré sur les statistiques produites par la DEPP, et notamment celles portant sur les élèves sortis du système scolaire sans diplôme ou qualification. L'objectif de résultat issu des travaux européens est remplacé par un diagnostic chiffré porté par la direction du ministère grâce à ses évaluations et travaux statistiques réguliers. Le remplacement des enjeux de la réforme dans un cadre national grâce à la mise en avant d'une sélection d'évaluations menées par la DEPP n'est sans doute pas étranger au contexte houleux qui préexiste à la ratification du nouveau Traité de Lisbonne. Ce temps télévisuel a en effet lieu avant la publication des résultats PISA 2003 et alors que les résultats du précédent cycle de l'évaluation ne sont pas réellement restés dans les mémoires des décideurs politiques et de l'opinion publique. Sans cette piqûre de rappel de l'existence d'évaluations internationales, le contexte dominant l'actualité en lien avec la politique européenne est celui des réticences à la ratification par la France du traité européen qui se font de plus en plus entendre, et avec la perspective du référendum qui sera organisé à l'été pour trancher la question. L'usage de PISA est alors porteur symbolique d'une intégration d'office de la réforme dans un champ qui n'est plus seulement celui de l'éducation mais qui est aussi – et presque surtout – économique et relatif au marché du travail futur. Et cela dans un contexte où cette dimension devient potentiellement conflictuelle, elle incite à se tourner vers les données nationales. Les données issues des travaux de la DEPP sont par ailleurs mieux appréciées des enseignants, qui y voient le regard de leur propre ministère de tutelle et non celui d'une organisation extérieure. Même si les syndicats d'enseignement n'ont pas encore exprimé de franche opposition à PISA durant ses cinq premières années d'existence, les données du ministère sont connues, ne serait-ce que parce que la DEPP mène depuis des années les évaluations-bilan sur lesquelles les organisations syndicales se sont régulièrement exprimées. A propos des arguments employés par le ministre lors de son intervention sur France 2, le SNUipp publiera d'ailleurs un communiqué de presse dans lequel il estime que l'intervention de François Fillon a été davantage calibrée pour « *ne pas heurter* » plus que pour développer en toute franchise les orientations de la réforme qu'il souhaite mener⁵⁸⁰.

⁵⁸⁰ SNUipp, « 100 minutes sans convaincre », Communiqué de presse du 19 novembre 2004.

Finalement, l'exposé des motifs du projet de loi ne mentionnera explicitement ni les évaluations nationales de la DEPP, ni PISA ou l'OCDE, même si les évaluations sont citées en filigrane dans le texte. Si les deux évaluateurs ont bien apporté leurs données et, pour l'OCDE, sa perception des enjeux du système éducatif à la lecture des données produites, leur influence dans la mise en politique publique des problèmes qu'ils identifient est encore faible et le seul nom des évaluations qu'elles mènent, notamment pour PISA, ne suffisent pas à asseoir une mise en politique publique des problèmes qu'elles identifient. Cette influence tient, pour la DEPP, principalement à la personne de Claude Thélot et à l'aura dont sa personnalité et sa conception de l'évaluation ont entouré le travail de la direction durant la décennie 1990. Pour l'OCDE, elle est encore très dépendante au début des années 2000 du contexte d'euphorie des questions éducatives, dans lequel elle a toutefois réussi à transférer les concepts clés sur lesquels elle travaille depuis plusieurs années, et notamment celui de l'acquisition de compétences comme résultat attendu de l'éducation, plus que l'acquisition de connaissances intégrées dans un programme scolaire. La sensibilité d'un gouvernement de droite au lien fait par l'OCDE et par la stratégie de Lisbonne entre enjeux de l'éducation et enjeux du marché du travail à venir n'est par ailleurs pas étrangère à la reprise par le gouvernement Raffarin du concept de compétences clés. Nous verrons dans l'étude du deuxième moment sur lequel nous nous arrêterons, la réforme portée par Vincent Peillon en 2013, que le rôle d'expertise de l'OCDE à travers PISA sera, moins de dix ans plus tard, décorrélé des clivages partisans à l'œuvre jusque-là et qui filtrent la réception des travaux de l'OCDE.

b) PISA, clé de voûte de la construction de la Refondation de l'école

De la même façon que la réforme de 2005 avait été initiée par une concertation citoyenne sur l'avenir de l'école, celle de 2013 commence avec une concertation menée à l'été 2012 en préalable à la rédaction du projet de loi qui sera présenté en conseil des

ministres et discuté en séance par la représentation nationale. Cette concertation, qui faisait partie des promesses de campagne du candidat Hollande et est considérée alors par les médias comme un dossier majeur de son quinquennat⁵⁸¹, est lancée le 5 juillet dans les murs de la Sorbonne, symbolisant l'approche guidant la démarche du ministre Peillon. L'environnement universitaire crée un détachement symbolique du travail préalable à la réforme du politique de l'image d'un projet conçu en chambre, en interne au cabinet et à l'administration centrale, et tend à le mettre en scène comme un projet travaillé sur le fond avec une rigueur proche du monde universitaire, et qui permettrait à l'École de « dépasser les querelles qui l'ont trop longtemps affaiblies » tout en lui offrant l'opportunité de « se tourner de nouveau vers l'avenir »⁵⁸². L'enjeu d'objectivation du débat est encore une fois mis en exergue, même si le principe est posé dès le départ de dépasser le niveau du constat pour aller vers la construction de propositions rationalisées. Bruno Julliard, conseiller de Vincent Peillon, déclare en conférence de presse « qu'il ne s'agit pas de faire un énième débat sur l'école qui commencerait par un état des lieux. L'état des lieux de l'école, il est connu, globalement il est même partagé »⁵⁸³. François Hollande, dans son discours de clôture, repose les deux évaluations qui donnent corps à ce constat de base. L'OCDE est nommément citée, et les conclusions de PISA sur l'accroissement des inégalités scolaires sont également utilisées pour évoquer plusieurs résultats d'évaluation qui sont menées par ou avec l'accompagnement de la DEPP, les inégalités territoriales de réussite aux examens de fin de collège selon les académies, mais aussi les évaluations de la maîtrise du socle de compétences attendu en fin de primaire, soit des années avant l'âge sur lequel porte PISA. Le Président de la République précise ainsi :

« L'école connaît aussi des échecs. Regardons la réalité en face. La France est l'un des pays où l'écart de résultats entre les élèves de statuts sociaux favorisés et défavorisés est le plus important des pays de l'OCDE. La réalité c'est que trop d'élèves, à la fin de l'école primaire ne maîtrisent pas les connaissances de base. La réalité c'est que les

⁵⁸¹ Le Monde, « Une concertation sur l'école très encadrée », 5 juillet 2012 ; Le Point, « Peillon et Ayrault lancent la concertation pour « refonder l'école », 5 juillet 2012.

⁵⁸² Alain Dulot, *Refondons l'école de la république*. Rapport de la concertation, 2013, p.7.

⁵⁸³ Le Point, *art. cit.*

inégalités territoriales se sont creusées le taux de réussite de l'Académie de Créteil est inférieur à celui de Nantes. »⁵⁸⁴

Des données sur le décrochage scolaire, encore une fois mises en regard comme en 2005 des conclusions de l'OCDE, sont immédiatement citées ensuite pour expliciter le second constat menant à la réforme.

« La réalité, c'est que trop de jeunes sortent du système éducatif sans diplôme ni qualification vous les avez chiffrés à 140 000, 140 000 dans une génération sans le passeport d'un diplôme, sans la fierté d'avoir obtenu un résultat au terme d'un parcours scolaire et qui connaissent l'échec, forcément l'échec, au lendemain de la sortie du système éducatif. »⁵⁸⁵

La production des données sur le nombre de décrocheurs n'est pas attribuée par François Hollande à la direction de l'évaluation du ministère, mais elle est associée aux travaux de la concertation, sans que leur source ne soit très clairement explicitée. Impossible de déterminer précisément la façon dont elles ont été obtenues, alors même que les enjeux qu'elles impliquent pour le Président de la République sont prioritaires à la fois pour les jeunes eux-mêmes et leur rapport à l'institution scolaire, et pour la société française et internationale des futurs actifs « *sans passeport* » qu'ils constitueront.

Le rapport de concertation d'une cinquantaine de pages qui est publié à la suite des travaux de la commission pour la refondation de l'école ne contient quant à lui ni les sources des documents utilisés pour fonder le constat chiffré du besoin de réformer en profondeur le système éducatif, ni les noms de l'intégralité des personnes qui ont participé à ses travaux. Seuls les membres du comité de pilotage sont précisés, tous connus soit pour être proches de la gauche soit pour leur connaissance fine des travaux d'évaluation menés par l'OCDE : François Bonneau, ancien conseiller d'orientation puis chef d'établissement avant de devenir conseiller régional du Centre-Val de Loire, Marie-Françoise Colombani, journaliste et grand reporter engagée dans la lutte pour le droit des femmes et le droit des personnes migrantes, Christian Forestier, haut

⁵⁸⁴ Discours de M. le Président de la République François Hollande à la suite de la concertation sur l'École, prononcé le 12 octobre 2012, p.2-3.

⁵⁸⁵ *Ibid.*, p.3.

fonctionnaire de l'Éducation nationale, enseignant chercheur et ayant occupé de nombreuses fonctions de haut rang au sein de l'Administration centrale du ministère, et Nathalie Mons, sociologue spécialiste de l'analyse des politiques éducatives et spécifiquement de PISA. Le tout est coordonné par Alain Dulot, romancier et agrégé de philosophie, qui en devient le rapporteur général.

Le rapport de la concertation⁵⁸⁶ cite largement dès ses premières pages les indicateurs de l'OCDE. Ils donnent une assise statistique à la première partie du texte qui interroge sur les raisons de la réforme souhaitée par le gouvernement, intitulée « *Pourquoi refonder l'école ?* ». Les indicateurs de l'OCDE sont employés pour tirer le constat des « *formidables atouts*⁵⁸⁷ » de l'école française, et pour donner les « *signaux* » qui « *doivent aujourd'hui nous alerter*⁵⁸⁸ ». A cet égard, les résultats des élèves à PISA sont repris aux côtés de ceux de l'évaluation de l'IEA PIRLS, qui évalue les élèves de primaire sur l'acquisition des compétences à l'écrit. Le constat tiré des résultats PISA depuis la première édition du programme est celui non seulement d'un positionnement moyen parmi les pays participants, dont la commission ne se satisfait pas, et d'une dégradation constante des résultats et notamment des plus fragiles. L'évaluation illustre également l'accroissement du mal-être à l'école, ou les inégalités de genre qui persistent encore en termes d'orientation scolaire. La présence au sein de la commission de Nathalie Mons, notoirement spécialiste de l'évaluation PISA et auteure de nombreux articles sur la réception de PISA en France, a pu jouer sur la façon dont l'évaluation a été intégrée aux travaux.

En revanche rien n'est écrit sur les résultats CEDRE, rien ne cite nommément les travaux menés par la DEPP. Seule une vague mention « *des enquêtes nationales* »⁵⁸⁹ est faite dans la première partie du rapport, après plusieurs pages de constats dressés à partir des données de l'OCDE et de PISA, ou illustrés par eux. Si vraisemblablement les indicateurs de la DEPP ont servi de base statistique à certains constats énoncés dans le rapport, comme celui de la prégnance des inégalités scolaires dès la maternelles

⁵⁸⁶ Comité de pilotage de la concertation sur la refondation de l'école, *Refondons l'école de la république*, 2012.

⁵⁸⁷ *Ibid.*, p.8.

⁵⁸⁸ *Ibid.* p.9

⁵⁸⁹ *Ibid.*, p.11.

et mise en évidence dans les évaluations faites en CM2, la direction ne tire aucune lumière de l'usage de ses travaux pour justifier le besoin de réforme sur ce problème identifié.

L'absence de citation des travaux utilisés dans le rapport officiel de la commission sur la refondation de l'école contribue à invisibiliser la DEPP dans la préparation du moment légitimant la loi telle que son histoire sera racontée par la gauche, alors même que les contributions des ministres qui portent la réforme dans les publications de la DEPP louent l'importance que son travail a joué dans la préparation du texte. Dans la préface du rapport annuel sur l'état de l'école, publié en octobre 2012⁵⁹⁰, les deux ministres signataires, Vincent Peillon et George Pau-Langevin⁵⁹¹, valorisent ainsi l'apport de la DEPP dans la préparation de leur réforme.

« Pour cela, nous avons besoin d'un diagnostic partagé. La concertation sur la Refondation a ainsi pu largement bénéficier des constats qui ont été établis, année après année, par l'état de l'Ecole. »⁵⁹²

Les travaux menés par la DEPP et synthétisés annuellement dans la publication *L'état de l'Ecole* créée par Claude Thélot lors de son passage à la DEPP, et dont la qualité est vantée par les deux ministres, sont toutefois caractérisés comme relevant du « constat » régulier, davantage que comme un outil de pilotage ou une source d'inspiration sur les problèmes à traiter ou les réformes possibles. Alors que PISA amène à s'interroger sur des problèmes publics et à y trouver des solutions à expérimenter, les travaux de la DEPP offrent une description d'un état de l'école année après année avec une antériorité intéressante mais peu problématisée et moins médiatique. Les évaluations nationales, dont celles de la DEPP, sont par ailleurs directement remises en cause par les auteurs du rapport sur la concertation. Dans deux chapitres intitulé de façon limpide « *Rénover l'action publique en éducation : des outils moins prescriptifs et plus efficaces* »⁵⁹³, et « *Une évaluation de l'Ecole recrédibilisée* », les auteurs questionnent non

⁵⁹⁰ DEPP, *L'état de l'école : coûts ; activités ; résultats : 31 indicateurs sur le système éducatif français*, n°22, octobre 2012.

⁵⁹¹ George Pau-Langevin est alors ministre déléguée à la réussite éducative.

⁵⁹² *Ibid.*

⁵⁹³ *Rapport de la concertation sur la refondation de l'école*, p.30.

seulement l'efficacité des outils déployés par la DEPP, mais aussi leur légitimité permettant de les utiliser comme guide pour les réformes à venir. Cette critique est d'autant plus centrale que l'évaluation est placée parmi les axes prioritaires à envisager pour une refondation de l'école réussie. La direction du ministère n'est, selon le rapport, pas suffisamment coordonnée avec les autres instances évaluatrices nationales pour offrir une vision cohérente et crédible scientifiquement de l'état du système éducatif et des enjeux nouveaux éventuels à prendre en compte pour l'avenir. Le défaut de crédibilité scientifique des évaluations nationales, incluant celles de la DEPP du fait de sa mission de service statistique ministériel, pourrait trouver une solution de remédiation en la création d'un processus de coordination des travaux existants par une entité indépendante.

« Plus généralement, nous souhaitons la mise en place d'une politique d'évaluation qui soit cohérente, légitime scientifiquement et participative. Pour cela, la création d'une instance indépendante qui coordonnerait cette politique d'évaluation, garantirait sa qualité scientifique et en diffuserait les résultats s'impose. »⁵⁹⁴

Plus directement encore, dans les propositions énoncées par les auteurs du rapport, la DEPP se voit mentionnée comme devant s'inscrire « *en complément des évaluations internationales*⁵⁹⁵ » dans le cadre de cycles d'évaluations des acquis des élèves français. Les auteurs regrettent par ailleurs que l'évaluation telle qu'elle est envisagée au niveau national ne permette pas d'avoir de « *vision globale du fonctionnement et de l'efficacité du système éducatif*⁵⁹⁶ », et qu'on ne dispose pas d'une instance évaluatrice qui assure « *la diffusion des résultats auprès des professionnels de l'éducation et des citoyens*⁵⁹⁷ ». Difficile de ne pas y voir une critique du fonctionnement de la DEPP, censée répondre à tous les manques pointés du doigt dans le rapport de concertation. Cette prise de position sur le besoin de revoir le fonctionnement de l'évaluation telle qu'elle est pratiquée par la direction du ministère contribue à expliquer pourquoi elle n'est pas citée nommément

⁵⁹⁴ *Ibid.*, p.30.

⁵⁹⁵ *Ibid.*, p.52.

⁵⁹⁶ *Ibid.*, p.52.

⁵⁹⁷ *Ibid.*, p.52.

au moment de l'utilisation de ses données pour asseoir les constats de départ de la concertation. Alors que, dans le cas de PISA, le nom même de l'évaluation accorde la validité scientifique nécessaire aux objectifs argumentatifs poursuivis, du fait de la reconnaissance commune de la valeur du programme de l'OCDE, le nom de la DEPP confère à ses données un désavantage majeur : celui de laisser craindre à une manipulation des chiffres par le ministre pour soutenir son propos, contre toute preuve scientifique.

Ce besoin de sortir l'évaluation du giron du ministère est également perçu, au moment du lancement de la concertation et des premiers échanges au sein de la majorité sur le contenu possible du projet de loi, comme un moyen encore une fois de sortir des oppositions classiques des acteurs des politiques éducatives, et d'éviter les blocages sans issue. Françoise Cartron, rapporteure socialiste de la loi au Sénat et proche de Vincent Peillon, explique ainsi en entretien le contexte de l'intégration des données de PISA à l'époque de la préparation du passage en séance.

« On sait former les très bons élèves. On sait former l'élite. On est là. On est très bien classé au classement international. Mais lorsqu'on regarde l'échec scolaire, là, on est très mal classé et bien sûr, avec une mission que ne remplit pas l'Éducation nationale, c'est à dire combler les inégalités. Bien sûr, les enfants les plus en échec sont les enfants qui sont issus des milieux défavorisés. Donc, ça pose un problème. Ça nous permet de regarder ce qui dépasse la vision purement Education nationale. C'est un vrai problème de société. Donc, ensuite, une fois qu'on a ces évaluations internationales, il s'agit de se retourner sur nos politiques et nos pratiques. Et de dire en quoi tous les efforts que l'on a pu faire ne résolvent pas ce qui est notre principal handicap. L'Education nationale avait peu l'habitude de faire ce type de réflexion. C'est pour ça que lorsqu'il y a eu la loi de refondation de l'école, je l'ai portée ici au Sénat avec le ministre Vincent Peillon. La loi, elle est partie du constat de l'OCDE et on a très souvent cité PISA en disant aujourd'hui, notre pays ne remplit pas bien sa mission qu'il doit remplir, ce qui avait un effet, ce qui nous fragilisait quelque part. Parce que les gens, les personnels de l'Education nationale se sentaient visés en disant « on dit qu'on ne fait pas bien notre travail, mais ce n'est pas nous. Qu'est ce que vous nous mettez entre guillemets sur le dos? »⁵⁹⁸

⁵⁹⁸ Entretien avec Françoise Cartron, Sénatrice socialiste de Gironde, Rapporteuse de la loi refondation de l'école, 10 mai 2017.

Le constat de PISA tel qu'il est formulé par Françoise Cartron se garde bien de désigner un responsable de la situation jugée problématique au vu des résultats de l'évaluation. Pour la sénatrice, l'objet de l'évaluation par PISA n'est finalement pas tant les élèves, ni même les enseignants, mais bien « *notre pays* ». Cet élargissement du champ pour désigner celui qui doit se réformer, ou du moins réformer ses priorités et ses pratiques, permet de déplacer l'objet de la réforme du corps enseignant vers un problème systémique et de société où tout le monde doit accepter le constat d'échec passé et tendre à réorienter son action dans le futur. En conséquence de ce dimensionnement du sens donné à PISA, les recommandations ne portent pas – ou pas uniquement - sur les pratiques des enseignants, mais bien sur ce qui relève de l'organisation de la scolarité et plus largement du lien entre école et société. Cette vision des conclusions permises par PISA introduit la problématique, centrale dans le projet de loi de Vincent Peillon, de la lutte contre les inégalités scolaires. Les données de l'OCDE ne sont pas les résultats des évaluations des résultats individuels des élèves, comme peut le faire la DEPP par exemple avec les évaluations exhaustives qu'elle mène tous les ans. Elles ne sont pas non plus le résultat d'une évaluation des pratiques enseignantes, qui se heurterait dans les solutions à apporter au principe de liberté pédagogique des enseignants inscrit dans le code de l'éducation. Elles donnent à voir par une évaluation (et une comparaison) internationale l'ampleur des dysfonctionnements à l'œuvre au sein du système éducatif français dont l'une des missions premières est de fournir à tous les élèves les conditions d'apprentissage leur permettant une pleine réussite scolaire. Par la mise en évidence de l'existence et de l'aggravation, plus que partout ailleurs, des inégalités scolaires en France, PISA montre le besoin d'une réorientation de la politique éducative française, d'autant plus que le thème de la reproduction scolaire des inégalités sociales est cher à la gauche (voir infra).

Les données et les conclusions de PISA sont clairement reprises dans l'exposé des motifs du projet de loi présenté en Conseil des ministres par Vincent Peillon le 23 janvier 2013, même si ni l'OCDE ni le nom de PISA n'est cité en toutes lettres.

« Le niveau global des compétences des élèves doit être amélioré pour parvenir à davantage de justice dans la réussite scolaire et pour inscrire le pays sur une trajectoire de croissance structurelle forte dans une économie de la connaissance internationale.

Depuis une dizaine d'années, le pourcentage d'élèves en difficulté face à l'écrit a augmenté de manière significative : près de 20 % des élèves de quinze ans connaissent de grandes difficultés de maîtrise de la langue écrite. Le nombre de sorties sans qualification, qui concernent 12 % des jeunes français d'aujourd'hui, doit être limité pour réduire le chômage qui touche en priorité ces populations. Enfin, le système éducatif français ne parvient pas à lutter suffisamment contre les déterminismes sociaux et territoriaux qui engendrent des inégalités sociales et géographiques et entraînent déclassement et crise de confiance pour une partie de la population. »⁵⁹⁹

Le ministre expose à plusieurs reprises l'enjeu de travailler à la lutte contre les inégalités scolaires au sens large, qu'elles soient liées à une origine socio-économique, géographique, un genre. Pour justifier l'enjeu, il s'appuie sur les résultats à une évaluation qu'il tient anonyme mais qui permet de comparer « *depuis une dizaine d'années* » (nous sommes en 2012, soit 11 ans après les publications des premiers résultats PISA) les performances à l'écrit des élèves de quinze ans. L'influence de PISA dans la construction de la légitimité de la réforme et dans la reconnaissance de sa capacité à montrer des problématiques mieux que quiconque d'autres est reconnue par le ministre et inscrite aux yeux de tous. La DEPP de son côté peine à faire reconnaître la valeur d'expertise de ses travaux, y compris de CEDRE, répondant pourtant aux besoins d'évaluation formulés par les membres de la commission pour la refondation de l'école, et spécifiquement celui de « *disposer en complément des évaluations internationales, d'évaluations régulières et suivies des acquis des élèves au niveau national, sur la*

⁵⁹⁹ Exposé des motifs du projet de loi sur la refondation de l'école.

base de la constitution d'échantillons »⁶⁰⁰ - c'est-à-dire exactement la méthode et la démarche employées par CEDRE.

2) L'usage de l'instrument dans le processus de construction de la mise en place de la volonté réformatrice

Le deuxième moment clé pour comprendre les modalités d'utilisation des évaluations statistiques sur les résultats des élèves et ce qu'elles disent de la nature de l'expertise des instances qui les produisent et les présentent est celui du débat législatif des projets de loi. L'observation des utilisations des évaluations nationales et internationales lors du passage de nos deux réformes par l'examen parlementaire donne à voir la façon dont elles sont mobilisées dans le cadre des deux objectifs classiques de légitimation de l'action de l'Etat et de constitution d'un contre-pouvoir à l'exécutif en demande de rationalisation des propositions formulées par le gouvernement⁶⁰¹. La comparaison temporelle et la mise en lumière des évolutions d'usage permettent ensuite d'envisager les transformations à l'œuvre du pilotage de l'action publique dans sa façon de matérialiser par des chiffres et des jugements d'experts les résultats qu'elle obtient par rapport aux objectifs qu'elle se fixe.

Nous observerons en particulier ici comment PISA et les travaux de la DEPP, dont en particulier CEDRE, ont été mobilisés, ou non, dans les deux moments de réforme, afin de voir leurs usages dans l'organisation de l'argumentation des débatteurs en séance. Nous nous appuierons sur deux interrogations relatives à la qualité de l'expertise dans le cadre de l'examen parlementaire d'un projet de loi. La première porte sur la reconnaissance publique du recours à un expert donné, comme légitimation de ses savoirs produits. Un expert qui n'est pas cité est-il toujours un expert ? En d'autres termes, qu'apporte le fait de citer le nom de l'évaluateur producteur de savoir par rapport au fait de ne pas le citer, et que cela signifie-t-il sur le

⁶⁰⁰ *Rapport de la concertation sur la refondation de l'école*, p.52.

⁶⁰¹ Clément Lacouette Fougère et Pierre Lascoumes. « L'évaluation : un marronnier de l'action gouvernementale ? », *Revue française d'administration publique*, vol. 148, n°4, 2013.

rôle d'expertise qui lui est – ou ne lui est pas – reconnu ? La seconde porte sur l'audition des évaluateurs producteurs de savoir. Au-delà de la mobilisation de leurs données dans les débats ou les productions écrites, la convocation aux auditions relève-t-elle un statut spécial de l'expertise de ceux qui ont la possibilité d'échanger avec les décideurs politiques sur leurs travaux ?

a) L'avenir de l'école : l'évaluation comme outil de certitude face aux enjeux européens

Le 8 février 2005, le ministre Fillon est auditionné au Sénat et à l'Assemblée nationale et y répète les mêmes constats au fondement de sa démarche de réforme. Il y fait valoir une exigence de « *lucidité* » face aux conclusions principales qu'il tire de diverses études, en particulier les évaluations internationales et notamment PISA. Le ministre précise ainsi :

« A l'évidence, le système éducatif a atteint depuis quelques années un palier qui ne permet plus à la France de progresser autant qu'elle le devrait: les performances de l'école française restent moyennes rapportées à celles des pays comparables ; sa contribution à la lutte contre les inégalités sociales s'est affaiblie ; son aptitude à garantir un niveau de formation qui réponde aux besoins de la société et de l'économie dans les prochaines décennies n'est plus assurée. »⁶⁰²

Plusieurs indices émanent de la déclaration introductive du ministre laissant apparaître le rôle de l'expertise de PISA et des travaux de l'OCDE en général dans le travail de formulation du problème public sur lequel porte la réforme. François Fillon s'appuie d'abord sur un constat « *d'évidence* », dressé par un évaluateur expert source de normativité décisionnelle externe⁶⁰³, c'est-à-dire un évaluateur d'étant pas décideur politique, et dont l'avis a conditionné les choix faits en matière de dispositif réformateur par le gouvernement. Il évoque ensuite les évaluations internationales et

⁶⁰² Audition de François Fillon à l'Assemblée nationale le 8 février 2005.

⁶⁰³ Pierre Lascoumes, « L'expertise, de la recherche d'une action rationnelle à la démocratisation des connaissances et des choix », *Revue française d'administration publique*, 3, n°103, 2002.

emploie les concepts clés de l'OCDE et son PISA, en pointant du doigt les « performances » de « la France », et non les résultats des élèves français par rapport à des « pays comparables » en termes d'enjeux sociétaux et économiques à venir. Les constats d'évidence sont donc supranationaux pour permettre la comparaison, ils portent sur des pays ayant des caractéristiques économiques proches comme peuvent l'être les pays européens et de l'OCDE, donnant des réponses aux questionnements de compétitivité future et de l'adéquation des formations dispensées à ces besoins posés dans le cadre de la stratégie de Lisbonne. Le concept de performance, absent du discours d'évaluation des politiques publiques et des résultats de l'éducation avant le début des années 2000, apparu dans le champ éducatif concomitamment au lancement de PISA sans que l'influence de l'OCDE pour sa popularisation ne soit très documentée, tend à laisser penser que PISA fait bien partie des inspirations du ministre. La mention des inégalités sociales, également un sujet d'intérêt de PISA, renforce encore cette lecture des sources de la réforme.

L'enjeu est à l'époque moins celui de démonstration de la performance de son système éducatif auprès de l'ensemble de la scène internationale, et donc de l'ensemble des pays participants, que de la compétition-coopération qui était lancée entre les pays européens au moment de l'adoption de la stratégie de Lisbonne. L'enjeu d'opérationnaliser le socle de compétence pour répondre rapidement et aux yeux de tous à la stratégie européenne rend la référence à l'OCDE intéressante pour le ministre car elle rentre dans les sujets potentiels d'échange entre les représentants des pays membres au sein de l'organisation internationale, au moment des rencontres régulières organisées pour la préparation des cycles PISA. Lors du premier jour de débat, François Fillon expose d'ailleurs le lien qui existe selon lui entre OCDE et Europe, en formulant par un raccourci le cadre international dans lequel se joue l'importance des résultats à PISA :

« L'effort sur les langues, la référence au cadre européen d'évaluation du niveau des élèves, l'apprentissage obligatoire d'une deuxième langue vivante, le socle de connaissances et de compétences fondamentales, tous ces thèmes, qui sont bien présents dans l'arrière-plan du texte qui vous est présenté, correspondent bien à des préconisations des instances européennes, de la Commission comme de l'OCDE.

La référence aux évaluations des résultats du système éducatif reste bien une des illustrations des difficultés de notre école. Les résultats des études PISA sont, vous le savez, moyens, mauvais en langues, plutôt meilleurs en mathématiques. Or notre rang dans la compétition européenne du savoir est bien l'une des préoccupations du Gouvernement. »⁶⁰⁴

L'inspiration de PISA et des travaux européens sur le socle commun sera d'ailleurs inscrite dans le marbre dans l'annexe intégrée dans le code de l'éducation une fois le projet de loi adopté en 2006.

« Le socle commun est le ciment de la Nation : il s'agit d'un ensemble de valeurs, de savoirs, de langages et de pratiques dont l'acquisition repose sur la mobilisation de l'école et qui suppose, de la part des élèves, des efforts et de la persévérance.

La définition du socle commun prend également appui sur la proposition de recommandation du Parlement européen et du Conseil de l'Union européenne en matière de "compétences-clés pour l'éducation et l'apprentissage tout au long de la vie".

Elle se réfère enfin aux évaluations internationales, notamment au Programme international pour le suivi des acquis des élèves (PISA) qui propose une mesure comparée des connaissances et des compétences nécessaires tout au long de la vie. »⁶⁰⁵

L'enjeu européen, adossé aux concepts clés de PISA de performance et la méthode de comparaison internationale typique de la façon de présenter les résultats de l'évaluation, est repris abondamment à la fois dans les rapports des commissions parlementaires et dans les débats, marquant le succès de l'orientation des échanges faite par le ministre. Dans son intervention liminaire en discussion générale, François Fillon mentionne clairement dès le départ l'impact d'un mauvais classement sur l'image de la France auprès de ses partenaires internationaux, dont il fait l'un des enjeux majeurs de sa réforme.

« Mesdames, messieurs les députés, vous avez entre vos mains le premier projet de loi sur l'éducation qui affirme aussi clairement l'ambition

⁶⁰⁴ François Fillon, 3^e séance du 15 février 2005.

⁶⁰⁵ Annexe intégrée à la section A du chapitre II du titre II du livre 1^{er} du code de l'éducation, par le décret n°2006-830 du 11 juillet 2006, signé par Dominique de Villepin, Premier ministre, Gilles de Robien, ministre de l'éducation nationale, de l'enseignement supérieur et de la recherche, et François Baroin ministre de l'outre-mer.

européenne de la France et de son système éducatif. (...) Au sein de l'Europe du XXI^e siècle, je ne puis situer la France qu'au premier rang. Or, toutes les comparaisons internationales montrent que notre pays obtient des résultats médiocres, en particulier dans le domaine de la maîtrise des langues vivantes. Cette situation n'est plus acceptable ! Elle est préjudiciable à la capacité de notre pays d'occuper pleinement sa place dans le monde. Elle limite les possibilités pour chaque Français de participer au développement des échanges internationaux dans les domaines culturels, scientifiques et économiques. »⁶⁰⁶

La comparaison internationale permise par PISA est déjà un outil argumentatif puissant, dont la portée est présentée comme dépassant largement le simple champ scolaire pour avoir un effet jusque dans les champs économiques et scientifiques mais aussi culturels.

Les rapports des commissions de l'Assemblée nationale et du Sénat se saisissent également de la portée européenne de la réforme et en reconnaissent l'importance, créant les conditions favorables d'une montée en influence de l'outil PISA dans le pilotage du système éducatif français. A l'Assemblée nationale, le rapporteur Frédéric Reiss⁶⁰⁷ fait bien mention en introduction de la nécessité d'inscrire le texte de loi dans la stratégie européenne de Lisbonne, et de l'obligation d'élargir le champ pour envisager l'école comme le pourvoyeur de la compétitivité économique future.

« La construction de l'Union européenne porteuse d'un modèle social et culturel spécifique, la compétition internationale et la globalisation de tous les aspects de la vie économique obligent à penser l'école hors du strict cadre national, ne serait-ce que pour comparer ses pratiques et ses résultats à ceux des autres nations. N'oublions pas la « stratégie de Lisbonne » qui fixe à l'Union Européenne de devenir l'économie de la connaissance la plus compétitive et la plus dynamique du monde. »⁶⁰⁸

⁶⁰⁶ François Fillon, 2^e séance du 15 février 2005.

⁶⁰⁷ Député Les Républicains du Bas Rhin depuis les élections de 2002.

⁶⁰⁸ Rapport n°2025 fait au nom de la commission des affaires culturelles, familiales et sociales sur le projet de loi d'orientation pour l'avenir de l'école, par M. Frédéric Reiss, député, enregistré à la Présidence de l'Assemblée nationale le 9 février 2005, p.9.

Le rapport du sénateur Jean-Claude Carle⁶⁰⁹ précise quant à lui l'intégration du projet de loi dans les enjeux d'internationalisation, et plus particulièrement d'europanisation, des problématiques scolaires.

« Par ailleurs, l'école de la République résonne des défis et des évolutions qui traversent notre pays, et interpellent l'ensemble de la Nation : le défi de la cohésion nationale et sociale, le défi de la société du savoir, le défi de l'ouverture de l'école sur l'Europe et sur son environnement extérieur. »⁶¹⁰

Cette internationalisation des données clés nourrissant le débat sur les raisons de la réforme, qui n'est pas remise en question dans les hémicycles lors de l'examen du texte, donne en filigrane à l'OCDE un statut d'œil évaluateur et de pourvoyeur de problèmes publics abordés dans la réforme. Néanmoins, le cadre européen est encore celui qui fait référence, et l'ampleur de l'étude PISA qui le dépasse pour s'intéresser aux pays de l'OCDE et au-delà n'est pas encore au centre de l'attention des décideurs publics. Par ailleurs, si plusieurs mentions explicites des travaux de l'OCDE, dont PISA, sont faites tout au long des débats sur des sujets clés comme l'évaluation de la performance des élèves, d'autres sujets de politique éducative que traite par ailleurs PISA et qui seront quelques années après très médiatisés (le redoublement, les inégalités scolaires, ..) sont abordés en s'appuyant sur les travaux de la DEPP. En 2005 encore, les deux évaluateurs sont convoqués pour examiner le projet de réforme, aussi bien par le ministre ou par les parlementaires. Ils acquièrent, par ces citations, un rôle d'expert contribuant en toute transparence à l'anticipation politique d'une crise du système éducatif à venir. Ici, l'OCDE comme la DEPP servent de référence explicite pour argumenter les positions tenues par les différents protagonistes de l'examen parlementaire de la réforme. Le type d'usage des évaluations nationales et internationales est par ailleurs encore relativement proche lorsqu'il s'agit d'entrer dans le détail des propositions faites dans les articles du projet de loi.

⁶⁰⁹ Sénateur Les Républicains de Haute-Savoie, élu de 1995 à 2018.

⁶¹⁰ Rapport n°234 fait au nom de la commission des affaires culturelles, sur le projet de loi d'orientation pour l'avenir de l'école, par M. Jean-Claude Carle, sénateur, 9 mars 2005, p.7.

Les travaux de la DEPP d'abord sont mentionnés à quelques reprises, pour étayer les conclusions écrites faites par les rapporteurs. Il s'agit principalement de données descriptives qui permettent de poser des constats sur l'inadéquation de l'organisation du système scolaire avec les enjeux perçus pour lui dans les années à venir. Frédéric Reiss mentionne par exemple les indicateurs portant sur le taux d'accès aux différents niveaux de formation avec un comparatif depuis les années 1980, ainsi que le taux d'accès au niveau IV en 2003 selon la filière et le sexe des élèves. La DEPP n'est pas sollicitée pour apporter son analyse des données. Elles sont problématisées par les rédacteurs du rapport de la commission et servent à appuyer l'argument du rapporteur du besoin de revaloriser les filières de l'enseignement professionnel et technologique qui souffrent toujours d'une image négative. Au-delà de la photographie à un instant « t » du système éducatif, les travaux de la DEPP sont également utilisés pour justifier l'importance de problèmes publics soulevés également par PISA, et notamment la reproduction sociale à l'œuvre à l'école. Le rapporteur Reiss souligne ainsi que la démocratisation de l'enseignement ne s'est pas accompagnée d'une « démocratisation des résultats⁶¹¹ ». Il précise que « le niveau d'études à la sortie du système éducatif est trop largement conditionné par l'origine sociale⁶¹² », et cite les travaux de la DEPP pour représenter les inégalités de parcours scolaire d'élèves ayant fait leur entrée dans l'enseignement secondaire en 1989. Si PISA aborde déjà la question des inégalités scolaires dans ses deux premiers rapports en 2001 et en 2004, c'est encore la direction du ministère qui bénéficie de la plus forte popularité ou de la plus forte légitimité auprès du rapporteur et de l'administration qui accompagne la rédaction de son rapport, et qui voit donc ses travaux cités.

Au Sénat, l'usage des données de la DEPP dans le cadre des travaux du rapporteur de la commission des affaires culturelles et de l'éducation Jean-Claude Carle est également double. Elles construisent d'une part un étayage à l'argumentation du rapporteur sur l'importance d'un enjeu de la politique scolaire. D'autre part, elles sont reprises pour les conclusions qu'elles tirent, ne nécessitant pas d'analyse

⁶¹¹ Rapport n°2025 fait au nom de la commission des affaires culturelles, familiales et sociales sur le projet de loi d'orientation pour l'avenir de l'école, par M. Frédéric Reiss, député, enregistré à la Présidence de l'Assemblée nationale le 9 février 2005, p.22.

⁶¹² *Ibid.*, p.23.

supplémentaire de la part du rapporteur. La direction est citée une fois dans le corps rapport de la commission, et deux de ses travaux sont référencés en note de bas de page pour parler des résultats des élèves français dans l'évaluation de leur maîtrise du langage et de la langue française à la fin de l'école primaire⁶¹³, et de l'effet de l'absentéisme des élèves sur la potentialité d'un décrochage scolaire chez les élèves de lycée⁶¹⁴. Jean-Claude Carle aborde également la problématique du redoublement et l'illustre par les données de la DEPP, venant appuyer un avis rendu par le HCée et faisant une comparaison internationale sur le recours au redoublement dans plusieurs pays européens comme les pays scandinaves ou le Royaume-Uni⁶¹⁵. Les données sont utilisées pour soutenir la thèse de l'utilité du redoublement lorsqu'il intervient tardivement dans la scolarité d'un élève, et notamment au lycée.

Les données de PISA sont également nommément utilisées par les rapporteurs, à la fois pour formuler des comparaisons internationales permettant de « révéler » que les performances des élèves français sont en réalité moyennes quand on les regarde au prisme des élèves des autres pays participants, et pour mettre en lumière des problématiques jugées prégnantes et inquiétantes, d'autant plus au regard de la comparaison avec les pays participants à PISA. Les rapporteurs n'hésitent pas à reprendre les données précises des résultats des élèves français sans que cela ne soit réellement signifiant pour un lecteur non rompu aux méthodes utilisées par l'OCDE pour son étude d'un nouveau genre.

« Les performances des jeunes Français se situent dans la moyenne des pays de l'OCDE. La seconde enquête PISA réalisée en 2003, qui évaluait principalement la culture mathématique de jeunes de 15 ans dans 40 pays, place la France en 16e position, avec un score légèrement supérieur à la moyenne des pays de l'OCDE (511 points contre 500). Le nombre d'élèves

⁶¹³ DEPP, « La maîtrise du langage et de la langue française à la fin de l'école primaire », *Note d'évaluation* n° 04.10, 2010.

⁶¹⁴ Étude de la Direction de l'évaluation et de la prospective (DEP) du ministère de l'éducation nationale conduite entre janvier et avril 2004.

⁶¹⁵ Haut conseil de l'évaluation de l'école, *Le redoublement permet-il de résoudre les difficultés rencontrées au cours de la scolarité obligatoire ?*, avis n°14, décembre 2004 ; à partir du rapport établi par Jean-Jacques Paul et Thierry Troncin.

présentant de graves difficultés de déchiffrage est passé de 4,2 % lors de la précédente enquête réalisée en 2001 à 6,3 %. »⁶¹⁶

Les références aux évaluations nationales et internationales sont en revanche bien moins assumées dans le cadre des deux fois cinq jours de débats oraux en février et mars 2005. Les travaux de la direction du ministère ne sont tout simplement pas utilisés, y compris pour aborder les taux de décrochage scolaire dont elle produit une partie des données avec l'INSEE. PISA, associé presque systématiquement au nom « OCDE », est quant à lui mobilisé à plusieurs reprises par le ministre et par les parlementaires. Le ministre d'abord affirme clairement aux yeux de tous le lien entre les recommandations de l'OCDE et les résultats de PISA, et le contenu de son projet de loi. Le soir du premier jour en séance à l'Assemblée nationale, il répond à l'intervention du député Liberti en précisant l'importance que le gouvernement accorde aux classements de la France dans les évaluations internationales, et cite même nommément non seulement les travaux de l'OCDE, mais aussi PISA.

« En effet, la réforme que nous proposons fait référence à l'Europe. L'effort sur les langues, la référence au cadre européen d'évaluation du niveau des élèves, l'apprentissage obligatoire d'une deuxième langue vivante, le socle de connaissances et de compétences fondamentales, tous ces thèmes, qui sont bien présents dans l'arrière-plan du texte qui vous est présenté, correspondent bien à des préconisations des instances européennes, de la Commission comme de l'OCDE.

La référence aux évaluations des résultats du système éducatif reste bien une des illustrations des difficultés de notre école. Les résultats des études PISA sont, vous le savez, moyens, mauvais en langues, plutôt meilleurs en mathématiques. Or notre rang dans la compétition européenne du savoir est bien l'une des préoccupations du Gouvernement. »⁶¹⁷

PISA sera ensuite cité principalement par des parlementaires de gauche au moment de défendre leurs amendements. Jean-François Voguet, sénateur du Val-de-Marne membre du groupe Communiste, Républicain et Citoyen, l'utilise pour donner en

⁶¹⁶ Rapport n°234 fait au nom de la commission des affaires culturelles, sur le projet de loi d'orientation pour l'avenir de l'école, par M. Jean-Claude Carle, sénateur, 9 mars 2005, p.13.

⁶¹⁷ 3^e séance, 15 février 2005.

exemple le suivi individuel des élèves en difficulté en Finlande, en Suède et au Japon et met en parallèle leurs bons résultats à PISA. La sénatrice d'Isère Annie David, du même groupe, utilise PISA pour mettre en lumière les difficultés des élèves français à mobiliser les connaissances acquises en classe dans des situations nouvelles. A l'Assemblée nationale, le député Yves Durand, qui sera en 2013 rapporteur du projet de loi pour la refondation de l'école de Vincent Peillon, est le seul à citer PISA deux fois au long des cinq jours d'examen du texte, pour appuyer ses amendements. A droite de l'échiquier politique, seul Gérard Longuet, sénateur Les Républicains de la Meuse, évoquera PISA pour alerter sur ce qu'il perçoit devenir un nouvel évaluateur, sur lequel les décideurs politiques nationaux n'ont pas réellement de prise.

La citation des travaux de l'OCDE témoigne de deux choses. D'abord, le nom de l'organisation et de son évaluation phare pour ce qui est des politiques éducatives s'inscrit dans une démarche de transparence des données utilisées pour argumenter la réforme. L'OCDE et PISA sont jugés externes aux travaux menés par le ministère et sont présentés comme des évaluateurs et évaluations indépendants, même si la DEPP est fortement impliquée dans la façon dont l'OCDE conduit ses évaluations du système éducatif en France. Ensuite, le nom de l'OCDE est employé pour gager de la crédibilité des données issues de PISA, plus que le nom de PISA lui-même qui n'a pas encore complètement acquis en 2005 la valeur d'un nom propre donnant corps à une entité à part entière. PISA est encore un programme de l'OCDE, et c'est bien le crédit accordé par un ministre de droite à l'organisation internationale qui est affirmé en premier lieu. L'OCDE donne légitimité à PISA, et non l'inverse.

La DEPP en revanche n'est pas à proprement parler citée lors des débats en séance, même si ses données sont employées et mises au crédit des travaux du ministère. Elle n'est pas non plus auditionnée par les commissions saisies sur le fond, pour expliquer ses travaux et répondre à d'éventuelles questions sur leur méthodologie ou les facteurs explicatifs des résultats obtenus. Cela interroge sur la nature de l'expertise accordée à la DEPP. La nécessité de transparence de l'expertise et de l'indépendance de l'expert du commanditaire inhérente à sa tâche n'est pas satisfaite suffisamment clairement dans le cas des travaux d'évaluation de la direction. Si elle

est jugée experte au sens où elle dispose d'une maîtrise technique des méthodes d'évaluation, de l'accès à des résultats exhaustifs sur les performances des élèves avec une antériorité, et dont d'une connaissance fine des résultats qu'elle apporte, son indépendance est loin d'être unanimement reconnue. L'absence de mention de l'origine des données de la DEPP entérine son intégration au ministère de l'Education nationale et le possible lien d'intérêt ou de subordination entre le ministre et « sa » direction de l'évaluation.

L'OCDE de son côté, malgré l'intérêt argumentatif de ses travaux et de PISA en particulier visible à travers leur utilisation dans les rapports des rapporteurs, n'est pas non plus auditionnée pour préparer les débats en séance et asseoir la position des commissions saisies sur le fond du projet de loi, ni au Sénat ni à l'Assemblée nationale. Cette non-convocation des experts pour rendre compte de ses travaux atténue considérablement la portée qui leur est accordée par les parlementaires. L'audition est en soi une reconnaissance non seulement de la qualité du travail d'un acteur et de sa pertinence pour éclairer les échanges sur les dispositifs prévus par le projet de loi. L'inscription de l'expertise produite, dans notre cas des indicateurs de résultats chiffrés, dans le cadre d'un échange entre décideurs publics et évaluateurs témoigne du fait que les données ont été lues et jugées pertinentes. Elle témoigne aussi au-delà de ça d'un besoin d'échange formulé par les décideurs politiques, afin d'une part de permettre à l'ensemble de la commission de pouvoir prendre connaissance des travaux d'expertise et interroger directement le producteur pour lever les incompréhensions éventuelles ou creuser les implications des conclusions produites ; et d'autre part d'aller plus loin que la simple lecture unidirectionnelle des chiffres et des recommandations générales pour tous les pays participants, et demander la parole de l'expert sur l'adéquation spécifique de ses jugements avec le cas particulier de la réforme en cours d'examen. Si la requête de l'éclairage resserré de l'expert OCDE sur la politique éducative française n'est pas formulée en 2005, signe d'une influence encore en construction de l'organisation internationale dans le cercle des experts en évaluation de la politique éducative française, elle deviendra centrale à la réforme qui sera conduite quelques années plus tard par Vincent Peillon.

b) L'évaluation dans les débats du projet loi de 2013: tous d'accord grâce à PISA ?

L'examen de la loi portée par Vincent Peillon donne d'emblée une toute autre portée aux données d'évaluation des résultats des élèves, et tout spécifiquement à celles produites dans le cadre de PISA. La communication faite par le ministre sur le fondement de sa réforme place PISA au centre du discours, à la fois pour les résultats qu'il fournit et pour la démarche de réflexion autour des orientations souhaitables du système éducatif impliquant une réforme de son fonctionnement et de ses objectifs.

i. PISA, et de droite et de gauche ?

L'ouverture de la séance à l'Assemblée nationale donne à voir dès les premières minutes l'influence de PISA dans le pilotage de la réforme de Vincent Peillon, à la fois pour justifier la démarche de réforme et pour fonder les mesures préconisées dans le projet de loi. Le ministre y donne en exemple, dès les premières secondes de son intervention, la réaction allemande suite aux premiers résultats PISA, ayant abouti à des réformes d'ampleur de leur système éducatif. La présence de la ministre allemande chargée des affaires culturelles, Annegret Kramp-Karrenbauer, venue à sa demande assister à l'ouverture du débat sur le projet de loi, lui permet de mettre en scène son acte réformateur devant un partenaire international clé, et de non seulement matérialiser l'exemple allemand devant la représentation nationale, mais aussi de personnifier la démarche de comparaison internationale utilisée par l'OCDE dans PISA⁶¹⁸. Au Sénat de la même façon, la mise en scène en moins, Vincent Peillon donne en exemple la démarche adoptée par « *les pays qui réussissent* », au premier rang desquels

⁶¹⁸ 1^{re} séance du lundi 11 mars 2013, intervention du ministre Vincent Peillon : « *Permettez-moi de saluer Mme Annegret Kramp-Karrenbauer, ministre plénipotentiaire de la République fédérale d'Allemagne, chargée des affaires culturelles dans le cadre du Traité de l'Élysée sur la coopération franco-allemande. Après l'avoir reçue ce matin, je veux la remercier chaleureusement d'avoir répondu favorablement à mon invitation à assister à l'ouverture de ce débat. C'est l'occasion pour moi de saluer la réussite d'un pays qui, après le choc provoqué par des résultats décevants au fameux test international PISA mené par l'OCDE, a réussi, au début de la décennie, à mettre en place des réformes ambitieuses qui lui ont permis d'améliorer les résultats de ses élèves. C'est ce que nous devons faire, c'est ce que nous allons faire* ».

il place l'Allemagne et ses actions prises pour réformer son système éducatif à la suite « du choc provoqué par les résultats aux tests PISA. »⁶¹⁹

Par cette revendication du lien entre PISA et la réforme qu'il veut mener, le ministre rend visible l'expert à l'origine de l'argumentation fondant son projet. L'usage de PISA inscrit encore une fois le débat dans le cadre d'enjeux internationaux attribués à l'éducation, devant les yeux du partenaire allemand. Il permet dans le même temps de construire des conditions favorables à l'adhésion par l'opposition aux propositions formulées au gré des articles du texte. PISA est en effet reconnu par les parlementaires de l'opposition pour ses qualités techniques et la légitimité de l'organisation qui le produit, issue de ses productions dans le champ des politiques économiques et conférant à ses résultats un statut de « certitude ». A l'opposé du mouvement de méfiance permanente à l'égard des jugements d'experts qui avait pu être étudié au début des années 2000⁶²⁰, l'évaluation internationale permet de réduire la possible controverse autour de l'incertitude du constat posé et de la manipulation politique des chiffres, et laisse espérer une adoption du projet de loi sans contestation majeure. L'évaluation PISA se voit donc officiellement légitimée par les parlementaires de droite comme de gauche, privant ainsi l'opposition, avec un « effet boomerang ⁶²¹» d'un de ses outils de défense à la fois de son bilan et de ses propositions d'amendement.

La légitimité transpartisane des données PISA se voit notamment au nombre de parlementaires qui les utilisent dans le débat pour prouver leurs propos, et à leur appartenance à une multiplicité de groupes politiques positionnés tout autour de l'hémicycle. En tout, sur les 30 interventions de sénateurs, 19 sont faites par des groupes de gauche. Elles ont lieu principalement avant mai 2012 et après mai 2017, lorsque les parlementaires se trouvent dans l'opposition au gouvernement en place. 19 sont faites par des groupes de droite, et une par un groupe du centre. A l'assemblée

⁶¹⁹ Séance du 21 mai 2013.

⁶²⁰ Voir par exemple Pierre Lascoumes, « L'expertise, de la recherche d'une action rationnelle à la démocratisation des connaissances et des choix », *Revue française d'administration publique*, 3, n°103, p. 2002 ; Philippe Roqueplo, « L'expertise scientifique, consensus ou conflit ? », *La terre outragée, les experts sont formels, Autrement*, série sciences en société, n°1, 1992, *Entre savoir et décision, l'expertise scientifique*, INRA, 1997.

⁶²¹ Xavier Pons, « PISA au Parlement ou l'effet boomerang d'un outil de communication politique », *Administration & Éducation*, vol. 145, n°1, 2015.

nationale, 44 des 113 interventions des députées sont faites par les groupes de gauche, 63 par les groupes de droite et 6 par les groupes du centre. Au-delà de ces répartitions relativement équilibrées par grandes tendances politiques, les thématiques abordées donnent à voir les indicateurs de PISA privilégiés par les différents groupes. Au Sénat, Jacques-Bernard Magner, du groupe socialiste, cite PISA pour insister sur le questionnement que doit selon lui entraîner les mauvais résultats et le mauvais classement de la France à PISA, tout comme Colette Mélot, du groupe UMP qui introduit son intervention en discussion générale par le constat de crise de l'école française rappelé par l'évaluation PISA. Jacques Legendre, du même groupe de droite, reprend un thème classique de sa famille politique pour prouver par une interprétation des indicateurs PISA la décorrélation qui existerait entre les moyens injectés dans le système éducatif et les bons résultats à l'évaluation. Un raisonnement similaire est repris par Sophie Primas lors de sa présentation des motifs du dépôt par le groupe UMP d'une question préalable. Marie-Christine Blandin, du groupe écologiste et présidente de la commission de la culture, prend la parole pour expliciter les motifs de son opposition à la motion de la droite, et cite elle-aussi PISA pour s'alarmer du « *verdict de l'étude, qui démontre que l'école en France est devenue la plus inégalitaire d'Europe*⁶²² », et introduit le sujet des inégalités scolaires comme l'un des apports centraux de PISA, dans la lignée des interventions du ministre. PISA est également largement mobilisé par les députés de tous groupes. Benoist Apparu, du groupe UMP, est le premier cette fois à évoquer le constat de la « *tendance* » du système éducatif à « *reproduire les inégalités de naissances*⁶²³ » et son aggravation au fil des évaluations PISA. Le même indicateur est repris quelques minutes plus tard par Barbara Pompili, du groupe écologiste, par Frédéric Reiss, député UMP et ancien rapporteur de la réforme Fillon, par Philippe Gomes du groupe UDI, et par Vincent Feltesse du groupe Socialistes républicains et citoyens. La députée UMP Virginie Duby-Muller aborde quant à elle la baisse de niveau qui s'aggrave d'évaluation PISA en évaluation PISA, dont la DEPP vient accroître la valeur puisque que son constat est « *confirmé par nos*

⁶²² Séance du 22 mai 2013.

⁶²³ 1^{ere} séance du 11 mars 2013.

*propres études nationales menées à l'issue des évaluations de CM2*⁶²⁴». Paul Salen, du même groupe de droite, utilise PISA non pas pour asseoir un jugement fondé sur l'état du système éducatif français, mais pour trouver les coupables de la situation qu'il juge alarmante. Le cycle 2009 permet alors de montrer que les mauvais résultats sont ceux des élèves ayant commencé leur scolarité en 1999 et 2000, soit sous le gouvernement Jospin et le ministère Lang.

On le voit à travers ces nombreuses illustrations, l'usage de PISA n'est pas l'apanage des parlementaires de droite, même si les indicateurs retenus par les uns et par les autres ne vont pas à contre-courant de leurs positions politiques traditionnelles. La construction de ce discours par la gauche ne date pas de leur arrivée à la tête du gouvernement français. On le voit déjà lors des débats en 2005, certains indicateurs de PISA trouvent grâce aux yeux des décideurs politiques de gauche qui n'hésitent pas à publiquement s'y référer. L'opportunité offerte par le travail de l'organisation internationale sur le poids des inégalités dans la réussite scolaire permet de lancer la réforme sans réserve sur le fond de l'évaluation, et sur le terrain des décideurs politiques de droite. La réponse du ministre à Jacques Legendre, sénateur du groupe UMP, qui interrogeait le gouvernement en février 2011, en séance de questions cribles thématiques sur l'aggravation des inégalités scolaires sur les conclusions à tirer des derniers résultats PISA, posait clairement le fondement argumentatif qui sera repris par la gauche au moment de la réforme de 2013 :

« La politique que nous menons va dans la bonne direction, non parce que c'est la nôtre, mais parce que c'est celle qui a été mise en place par tous les pays obtenant de meilleurs résultats que nous dans les enquêtes PISA. »⁶²⁵

Cet aplatissage des oppositions politiques aux données produites par PISA est également constaté en interne par les équipes de l'OCDE, qui reconnaissent avec les débats de 2013 l'alignement des paramètres politiques et sociétaux renforçant la légitimité de leur évaluation dans le pilotage de la politique publique nationale. Le

⁶²⁴ 1^{re} séance du 12 mars 2013.

⁶²⁵ Réponse du ministre de l'éducation nationale à la question crible thématique n°0153C de Jacques Legendre, UMP, publiée au journal officiel du Sénat le 9 février 2011.

porte-parole de la direction de l'Éducation de l'OCDE en France, Eric Charbonnier, reconnaissait ainsi en entretien ce nouveau positionnement politique des données PISA, permettant de fermer au nez des parlementaires de l'opposition les fenêtres de critique du bienfondé de la réforme. L'OCDE permet d'étaler aux yeux de tous ce qui relève du bilan du gouvernement précédent, et « il ne peut pas y avoir de discussion ».

« Aujourd'hui il y a une volonté d'utiliser des résultats pour pousser des réformes. Mais c'est aussi plus facile. Quand on a sorti les résultats de 2012, automatiquement c'était le bilan du gouvernement précédent. Cette fois-ci il ne pouvait même pas y avoir de discussion, c'est plus facile dans ce cadre-là d'utiliser les résultats pour pousser les réformes. Je ne suis pas sûre que ça soit lié au parti politique mais il y a quand même une volonté forte de s'attaquer aux inégalités, de donner la priorité au primaire, des choses qui sont un peu nouvelles et plus en accord avec les recommandations qu'on faisait depuis une dizaine d'années. » ⁶²⁶

L'arrivée d'un gouvernement de gauche, historiquement attaché à lutter contre les inégalités scolaires dès les premières années de scolarité, permet de valoriser des indicateurs produits par PISA et qui n'avaient pas encore réellement et l'écho massif souhaité par l'OCDE.

Les usages des indicateurs dans la construction de la réforme, finalement en ligne avec les orientations politiques prônées de longue date par les deux groupes politiques même s'ils ne rebattent pas les cartes de la rhétorique politique⁶²⁷, montrent bien que PISA n'est plus considéré comme un outil teinté des idées néolibérales de l'OCDE mais bien comme une source d'expertise tout à fait légitime pour évaluer les résultats du système éducatif français quel que soit son parti politique. En construisant des indicateurs portant à la fois sur les effets du financement du système éducatif et sur les effets de la construction sociale des pays sur les résultats des élèves, l'OCDE a créé un outil malléable et polysémique abordant les thématiques clés d'un panel large de formations politiques, renforçant son attrait argumentatif et de communication politique. La publication des résultats PISA en 2010 et à la fin de l'année 2013,

⁶²⁶ Entretien avec Eric Charbonnier, Analyse à la Direction de l'éducation, OCDE, 16 décembre 2016.

⁶²⁷ Xavier Pons, « PISA au Parlement ou l'effet boomerang d'un outil de communication politique », *Administration & Éducation*, vol. 145, n°1, 2015.

confortent les orientations souhaitées par le gouvernement et incitent à les utiliser sur un temps long, à la fois au moment de la réforme et à chaque nouvelle publication, pour insister en continu sur le bienfondé de l'action ministérielle Agathe Cagé, ancienne conseillère de Vincent Peillon, expliquait en entretien l'utilité récurrente de PISA pour appuyer les idées portées par un ministre de gauche, et le changement de regard de la part des décideurs politiques de gauche suite au constat de l'adéquation de leur volonté réformatrice et des résultats fournis par l'évaluation de l'OCDE.

« Depuis 2012, on a un regard positif sur PISA. On s'en sert de guide. Pour la loi refondation de l'école il y a un guide annexé qui fait référence à PISA. Les résultats de PISA 2012 en décembre 2013, quelques mois après le début de la loi refondation de l'école, nous ont fait dire que ça allait dans le bon sens. PISA 2012 ça a été un choc, c'est une claque dans le sens que ça va dans le sens du constat qu'on fait, dans le sens de la politique conduite depuis quelques mois. C'est une vraie opportunité. »⁶²⁸

Le succès de la popularisation et la légitimation politique de PISA est d'autant plus grand que son ambition de lier sa vision des enjeux de différents champs de politiques publiques, et notamment la politique éducative et la politique de l'emploi, porte ses fruits y compris auprès des parlementaires de gauche. L'auto-référence aux conclusions de PISA dans l'ensemble de ses produits dérivés et dans les travaux transversaux à plusieurs directions de l'OCDE a fonctionné au moins pour ce qui est du partage des objectifs de l'éducation pour le marché du travail, et la lecture de l'avenir économique des pays de l'OCDE que fournit PISA a, à force de répétition, été finalement récupérée par les décideurs politiques français et reversée aux travaux préparatoires à l'examen parlementaire de la réforme de 2013. Yves Durand, auteur de plusieurs rapports parlementaires sur les évaluations des politiques éducatives et la qualité de l'enseignement, rapporteur de la loi de 2013, est exactement dans la lignée de la conception portée par l'OCDE des objectifs de l'éducation lorsqu'il donne en entretien sa vision de ce que fait l'école.

⁶²⁸ Entretien avec Agathe Cagé, Conseillère au cabinet du ministre de l'Education Nationale en charge du second degré, des programmes et de l'évaluation des élèves, octobre 2016.

« Qu'est ce que fait l'école? On s'est jamais dit ça, mais l'école forme à des métiers et à une vie. Il y a des métiers dont on ne connaîtra pas les contours dans les dix ans qui viennent. On a parlé de l'école, mais on n'a pas suffisamment parlé de l'enseignement professionnel, qui est un peu le parent pauvre de toutes les réformes de l'Éducation nationale, qui devrait former à des métiers qu'on connaît pas encore. Ce n'est pas des métiers qu'il faut former, il y a des compétences qui peuvent permettre à des jeunes, des gens, de changer de métier. Et donc là on a le lien direct avec ce que fait l'OCDE. »⁶²⁹

La liste des auditions menées pour préparer les rapports des commissions donne à voir par ailleurs l'influence des évaluateurs dans la fabrique politique de la loi. Au Sénat, tout comme à l'Assemblée, Eric Charbonnier est auditionné pour l'OCDE par la commission. A la lumière des entretiens que nous avons réalisés avec les deux rapporteurs de la réforme, on s'aperçoit de la prépondérance de PISA dans la préparation de l'examen de la loi. L'OCDE est incontournable pour les deux rapporteurs, alors que les travaux de la DEPP, même s'ils les ont à l'esprit, ne sont pas réellement en haut de la liste des expertises qui comptent. Yves Durand, interrogé sur les experts auditionnés à sa demande par la commission de l'Assemblée nationale, cite immédiatement l'OCDE et les visites à l'étranger qui ont été organisées pour préparer les débats, afin de mieux saisir la portée des conclusions fournies par PISA.

« Quels experts avez vous auditionné pour préparer votre rapport ?

Dans ce cas précis (la préparation des travaux sur la loi de 2013), on avait bien sûr auditionné l'OCDE, et on avait été dans les pays étrangers, pour le côté international, aux Pays-Bas.

(Je relance) Et au niveau des instances nationales ?

On avait interrogé aussi la Dgesco, mais ce ne sont pas des bureaux d'évaluation.

Et d'autres évaluateurs ?

⁶²⁹ Entretien avec Yves Durand, Sénateur socialiste puis LREM de Mayenne, rapporteur de la loi refondation de l'école, président du comité de suivi et membre du comité d'évaluation des politiques publiques, 14 juin 2017.

La DEPP. On avait interrogé la DEPP. Je ne sais pas si on n'avait pas demandé à la Cour des comptes aussi, sur le volet financier. »⁶³⁰

La DEPP n'est pas spontanément citée, et il faudra deux relances pour que sa contribution au travail préparatoire de la commission revienne à l'esprit du rapporteur. Cette hiérarchie des évaluateurs, défavorable à la DEPP, transparaîtra dans la rédaction du rapport d'Yves Durand. Les résultats de PISA et en particulier son indicateur portant sur les inégalités scolaires apparaissent dès les premières lignes du rapport de la commission de l'Assemblée nationale.

« Particulièrement bien attestée sur le plan statistique, la frontière qui sépare les élèves pour qui l'école française est l'une des meilleures au monde de ceux pour qui elle est l'une des plus mauvaises conforte les inégalités sociales et nourrit les sorties sans diplôme ou qualification du système éducatif.

Ainsi, tous les indicateurs sont, objectivement, « au rouge ». En effet, le système éducatif se caractérise, depuis au moins dix ans, par une forme d'inertie dans la dégradation, tandis que les résultats de nos élèves sont, parmi tous les pays de l'OCDE, ceux les plus corrélés à l'origine sociale. »⁶³¹

Yves Durand insiste sur la valeur de l'expertise de l'OCDE, qui fournit des indicateurs PISA non seulement « objectifs » dénués de tout intérêt politique ou économique, mais aussi d'une qualité méthodologique reconnue unanimement. L'évaluation est utilisée comme argument clé pour mettre en avant le creusement des écarts de résultats des élèves selon leur milieu socio-économique, mais aussi pour insister sur le poids de toutes les autres inégalités qui sont étudiées par PISA et les réorientations de politique éducative qu'elles impliquent comme la lutte contre la fracture numérique à l'école,

⁶³⁰ Entretien avec Yves Durand, Sénateur socialiste puis LREM de Mayenne, rapporteur de la loi refondation de l'école, président du comité de suivi et membre du comité d'évaluation des politiques publiques, 14 juin 2017.

⁶³¹ Rapport n°767 fait au nom de la commission des affaires culturelles et de l'éducation *sur le projet de loi d'orientation et de programmation pour la refondation de l'école de la République, Tome 1*, par Yves Durand, député, enregistré à la Présidence de l'Assemblée nationale le 28 février 2013.

ou la réaffirmation de la pertinence du collège unique pour réduire les inégalités scolaires et améliorer la performance de l'ensemble des élèves.

Du côté des éléments d'expertise fournis par la DEPP, certains de ses travaux et en particulier CEDRE sont employés pour asseoir les constats faits par les rapporteurs à l'Assemblée comme au Sénat. Loin de constituer la source prioritaire de savoirs, ses évaluations servent souvent de compensation à l'absence de données PISA en date des échanges, ou de preuve supplémentaire de la conclusion énoncée. Le ministre Peillon le dit lui-même en filigrane lorsqu'il décrit à l'Assemblée nationale les sources de ses constats, qu'il ordonne de la façon suivante : d'abord le programme PISA, puis le rapport PIRLS, et « *enfin, les évaluations de la DEPP elle-même*⁶³² ». Tout en reconnaissant le rôle de la DEPP dans la mise en œuvre des évaluations internationales qu'il cite, il positionne tout de même ses travaux au dernier rang des expertises employées pour justifier des mesures de la réforme. Les références à CEDRE dans les rapports des rapporteurs sont souvent accolées à des résultats de PISA, et viennent renforcer les constats faits par l'OCDE. CEDRE est par exemple cité en introduction du rapport du Sénat, mais uniquement pour pallier le délai d'analyse des résultats de PISA 2012⁶³³. L'évaluation de la DEPP voit toutefois sa méthodologie rapprochée de celle de l'OCDE, témoignant de sa solidité et de sa légitimité.

« Depuis 2003, le ministère de l'éducation nationale a mis en place un dispositif rigoureux, le Cycle des évaluations disciplinaires réalisées sur échantillon (CEDRE). Renouvelées tous les six ans pour chaque grand domaine de compétences (maîtrise de la langue, acquis en langues vivantes, histoire-géographie, sciences expérimentales, mathématiques, attitudes à l'égard de la vie en société), ces évaluations se déroulent selon des modalités analogues à celles des enquêtes internationales comme PISA. »⁶³⁴

⁶³² 2^e séance du 13 mars 2013.

⁶³³ Rapport n°568 fait au nom de la commission de la culture, de l'éducation et de la communication sur le projet de loi d'orientation et de programmation pour la refondation de l'école de la République, par Françoise Cartron, sénatrice, enregistré à la Présidence du Sénat le 14 mai 2013, p.11.

⁶³⁴ *Ibid.*, p.12.

La stratégie de la DEPP pour valoriser ses travaux en profitant de l'aura de PISA semble en partie réussie, même si elle a pour corollaire de minimiser la portée explicative autonome des évaluations de la direction. La DEPP est citée plusieurs fois dans le rapport du Sénat, pour les données descriptives du système éducatif qu'elle propose et son analyse des évaluations internationales des résultats des élèves. Elle est citée deux fois en note de bas de page renvoyant à des notes d'information qu'elle avait rédigé sur l'évaluation menée par l'IEA sur les compétences des élèves de CM2 en lecture PIRLS et les résultats du baccalauréat 2012. Ses données sur l'évolution des effectifs scolaires sont également utilisées. Dans le rapport de l'Assemblée nationale, son évaluation CEDRE apporte des données chiffrées appuyant le constat du rapporteur de la faiblesse de l'enseignement des langues, le français écrit comme les langues étrangères en général.

Cette difficulté rencontrée par la DEPP à faire valoir la portée potentielle de l'expertise fournies par ses évaluations nationales est particulièrement visible lorsqu'on regarde les données mises en avant par le ministre Vincent Peillon au moment de montrer la pertinence de sa réforme pour redonner à l'école son identité républicaine et lutter contre l'aggravation des inégalités scolaires. PISA est cité d'abord puis plus discrètement CEDRE, dont les indicateurs portant sur ce sujet sont pourtant proches de ceux de l'OCDE. Le problème des inégalités scolaires en particulier, pierre angulaire de la réforme, sera mis à l'agenda politique publiquement à partir des données PISA. Cette catalyse de la réforme autour de PISA, qui légitime l'action réformatrice, lui attribue en retour un rôle d'expert sans contre-expert, ramenant la DEPP à sa fonction d'opératrice nationale de PISA.

ii. Le problème public des inégalités scolaires comme catalyseur de l'usage de PISA dans la réforme

Le projet de loi de Vincent Peillon place en son centre la lutte contre les inégalités scolaires, et met en exergue PISA pour justifier l'importance et l'urgence du problème. Les indicateurs de l'OCDE portant sur le lien entre performance et l'origine des élèves

sont directement nommés dans l'annexe du projet de loi, avant les résultats des évaluations nationales pouvant pourtant apporter un éclairage statistique à la démarche de réforme voulue par le ministre.

« Le niveau global des compétences des élèves formés en France doit être amélioré pour parvenir à davantage de justice dans la réussite scolaire et pour pouvoir inscrire le pays sur une trajectoire de croissance structurelle forte dans une économie de la connaissance internationale.

(...)

La France se classe dans les derniers rangs des pays de l'OCDE (27e sur 34 pays) du point de vue de l'équité scolaire, ce qui signifie que l'incidence de l'appartenance sociale sur les résultats scolaires y est plus forte que dans d'autres pays de l'OCDE. Les données statistiques nationales montrent l'importance et la persistance des écarts entre résultats scolaires selon les lieux de scolarisation et donc la difficulté de lutter contre les inégalités sociales : le pourcentage des élèves n'ayant pas atteint des acquis suffisants en français en CE1 est deux fois plus élevé dans certaines académies que dans d'autres. De même, le taux de réussite au baccalauréat général peut varier de près de dix points entre académies de la métropole, l'écart étant encore plus fort avec les académies d'outre-mer. Enfin, la maîtrise des compétences de base en 3e entre 2007 et 2011 s'est dégradée significativement pour les élèves de l'éducation prioritaire.

Ces inégalités mettent à mal la promesse républicaine, qui est de permettre la réussite de tous. La refondation doit conduire à une réduction de l'impact des déterminismes sociaux et de toutes les inégalités et les discriminations. »⁶³⁵

L'emploi de l'argument PISA avant l'argument des évaluations nationales témoigne de l'influence de premier ordre de l'évaluation de l'OCDE dans l'identification et la mise en politique publique des problèmes qu'elle soulève avec PISA. Cette reconnaissance de l'apport d'expertise supérieur de l'OCDE sur le sujet des inégalités scolaires, qu'elles soient socio-économiques, géographiques ou encore liées au genre des élèves, est finalement assez récente dans le débat public français. Si l'on reprend notre moment réformateur de comparaison qu'est la loi de 2005, alors que PISA a déjà donné à voir le positionnement de la France en termes de force du lien entre

⁶³⁵ Annexe à la loi n°2013-595 du 8 juillet 2013 d'orientation et de programmation pour la refondation de l'école de la République.

performance des élèves, l'OCDE est loin de tenir le rang d'expert privilégié, y compris dans l'argumentaire des parlementaires de droite. Ce problème précis avait pourtant été considéré comme un enjeu majeur dans le rapport Thélot suite au grand débat national, qui avait développé dans son point 4 la nécessité de réduire les inégalités scolaires en assurant notamment une meilleure répartition des moyens entre les établissements scolaires.

« L'École doit favoriser la mixité sociale et mobiliser l'ensemble de ses moyens dans une lutte contre les ségrégations de toutes sortes. (...) Pour tendre vers l'égalité des établissements, il faudrait mettre en place, à partir d'une évaluation publique renforcée, une politique plus ambitieuse qu'aujourd'hui de différenciation maîtrisée, c'est-à-dire de réduction volontariste des inégalités, en allouant des moyens nettement accrus à ceux qui sont confrontés à de réels problèmes tenant à leur environnement. »⁶³⁶

La question des inégalités scolaires apparaissait également dans les débats parlementaires, à la fois dans les rapports produits par les commissions saisies sur le fond et dans les discussions en séance. PISA n'était alors pas cité pour apporter des arguments chiffrés posant le constat d'un problème français de la persistance voire l'aggravation du problème des inégalités scolaires, ni pour légitimer les propositions pour y remédier. D'autres instances évaluatrices avaient été à l'époque préférée à l'organisation internationale – et à la DEPP. Le rapport de Jean-Claude Carle citait par exemple plusieurs études et avis rendus par le HCÉE, mettant en évidence le fait que « *la pauvreté est la première cause du grand échec scolaire* »⁶³⁷. Il reprenait également les conclusions du rapport du Conseil de l'emploi, des revenus et de la cohésion sociale (CERC) sur « *Les enfants pauvres en France* »⁶³⁸, pour ajouter que « *près de la moitié d'entre eux connaissent des difficultés scolaires importantes dès le primaire* »⁶³⁹. Les données présentées dans un graphique provenaient quant à elles de deux enquêtes de l'INSEE, Formation et qualification professionnelle et Enquêtes emploi, pour montrer le taux d'obtention

⁶³⁶ Rapport de la Commission du débat national sur l'avenir de l'École présidée par Claude Thélot, *Pour la Réussite de tous les élèves*, La Documentation Française, Paris, 2004, Synthèse IV.

⁶³⁷ Rapport n°234 fait au nom de la commission des affaires culturelles, *sur le projet de loi d'orientation pour l'avenir de l'école*, par M. Jean-Claude Carle, sénateur, 9 mars 2005, p. 14.

⁶³⁸ *Ibid.* p.14.

⁶³⁹ *Ibid.* p.14.

du baccalauréat selon la génération et le milieu social des élèves. Les données et les conclusions de PISA, citées pourtant quelques paragraphes plus haut pour montrer la baisse de performance des élèves français, n'apparaissent pas pour illustrer la prégnance du sujet des inégalités scolaires, encore traitées dans un cadre national en lien avec les problématiques de décrochage et les défauts d'une orientation forcée des élèves les plus défavorisés vers des filières ne correspondant pas à leurs aspirations professionnelles. Le lien direct entre données PISA et enjeu de réduire les inégalités scolaires n'était pas non plus fait en séance par les parlementaires, y compris ceux sensibles aux travaux de l'OCDE. Yves Durand, qui citait PISA déjà à l'époque pour insister sur l'importance des TPE supprimés par le ministre sans concertation des parlementaires, interpellait également sur l'absence de lien fait entre inégalité sociale et échec scolaire⁶⁴⁰. PISA n'était pas utilisé pour appuyer le lien de causalité mis en avant par le rapporteur.

Ce rôle d'expert clé, nouveau pour PISA et d'autant plus parce qu'il est utilisé par un gouvernement de gauche habituellement précautionneux voire réservé quant aux recommandations de l'OCDE en termes de politiques publiques, est lié à une conjonction de paramètres qui constituent ensemble un contexte favorable à la mise à l'agenda politique de ses sujets de travail et des préconisations qui en découlent. On retrouve ici les dimensions de la typologie formulée dans les années 1980 par John Kingdon⁶⁴¹ sur laquelle s'appuie Erik Neveu⁶⁴² pour étudier les conditions de mise en politiques publiques de problèmes identifiés par des entrepreneurs de politique publique, puisque trois conditions sont réunies pour faire de PISA l'expertise majeure de la réforme Peillon.

Premièrement, le problème de l'aggravation des inégalités scolaires est bien positionné au moment de la réforme en haut des priorités des décideurs politiques. Le thème est cher à la gauche et présent régulièrement dans les discours de ses représentants politiques et les réformes du système éducatif engagées lors de leurs périodes au pouvoir. Le problème public des inégalités scolaires a déjà émergé depuis

⁶⁴⁰ 2^e séance du 14 février 2005.

⁶⁴¹ John Kingdon, *Agendas, Alternatives and Public Policies*, New-York, Harper Collins, 1984.

⁶⁴² Erik Neveu, « L'analyse des problèmes publics. Un champ d'étude interdisciplinaire au cœur des enjeux sociaux présents », *Idées économiques et sociales*, vol. 190, n°4, 2017.

longtemps, et a déjà été traité à de nombreuses reprises par différentes réformes d'ampleur du système éducatif menées par la gauche, en 1981 et 1989 en particulier avec la mise en place d'une série de dispositifs nouveaux et de recentrage conceptuels chargés de réaliser l'idéal républicain de l'école ascenseur social, de la mise en place des zones d'éducation prioritaire au paradigme du placement de l'élève au centre du système éducatif⁶⁴³. Les inégalités scolaires, ou la capacité du système scolaire français à reproduire des inégalités sont par ailleurs étudiées depuis un demi-siècle et popularisées par des grands noms de la sociologie, au premier rang desquels Pierre Bourdieu et Jean-Claude Passeron⁶⁴⁴, ou Christian Baudelot et Roger Establet⁶⁴⁵. Les deux derniers, auteurs depuis le début des années 1970 d'ouvrages phares sur le sujet, ont d'ailleurs régulièrement écrit sur l'évaluation PISA et son inintérêt lorsqu'elle est abordée sous l'angle du palmarès⁶⁴⁶.

A la suite de ce constat d'antériorité on ne peut plus publique, il est difficile de considérer PISA comme l'outil ayant permis de « révéler » à proprement parler le problème public des inégalités scolaires en France. L'OCDE sort donc ici de la définition classique de l'entrepreneur de problème public au sens où l'organisation ne fait pas émerger un problème. Pourtant, il faut noter que le champ lexical qui entoure l'indicateur de PISA faisant le lien entre origine socio-économique et performance scolaire est bien souvent celui de la révélation. Isabelle Robin, responsable du département de la recherche et du développement, de l'innovation et de

⁶⁴³ Voir plus particulièrement la réforme lancée par Alain Savary en 1981 qui déploie les zones d'éducation prioritaire pour lutter contre l'inégalité des chances à l'école, et la réforme de 1989 lancée par Lionel Jospin qui verra l'élève placé au centre du système éducatif, considéré pour ses caractéristiques et besoins propres.

⁶⁴⁴ Pierre Bourdieu, Jean-Claude Passeron, *La Reproduction*, Paris, Editions de Minuit, Pierre Bourdieu, Jean-Claude Passeron, *Les Héritiers, les étudiants et la culture*, Paris, Editions de Minuit, 1964.

⁶⁴⁵ Christian Baudelot, Roger Establet, *L'école capitaliste en France*, Paris, Maspero, 1971 ; *L'école primaire divisée*, Paris, Maspero, 1979 ; *Le niveau monte*, Paris, Le Seuil, 1988 ; *Allez les filles*, Paris, Le Seuil, 1992 ; *Quoi de neuf chez les filles ? Entre stéréotypes et libertés*, Nathan, coll. L'enfance en questions, 2007 ; *L'élitisme républicain. L'école française à l'épreuve des comparaisons internationales*, Seuil, coll. La république des idées, 2009.

⁶⁴⁶ La critique de PISA qu'ils mènent en particulier dans leur ouvrage de 2009, l'élitisme républicain. *L'école française à l'épreuve des comparaisons internationales*, est fondé sur une analyse systématique des données de la France à PISA, permettant de discuter, après avoir démontré l'inutilité du raisonnement en palmarès, les idées reçues à propos du système éducatif français, des bienfaits du redoublement au lien entre égalité sociale et réussite scolaire en passant par l'effet de la proportion d'élèves enfants d'immigrés et une baisse générale des performances.

l'expérimentation à la Dgesco, revenait en entretien sur la réception des constats de l'OCDE à la sortie des cycles PISA 2006 et 2009 :

« PISA ça nous a parlé des inégalités scolaires, qu'on ne voyait pas ! Qu'on ne voyait pas ! L'école française s'est longtemps vécue comme égalitaire, Elle s'est longtemps vécue, s'est longtemps pensée, comme étant le modèle républicain, méritocratique, et là derrière vous avez toute l'image de l'élève boursier méritant qui accède aux plus hauts postes, de responsabilités de l'Etat etc etc. Donc sans PISA en fait, et sans les analyses mettant en regard les résultats et l'origine sociale on serait certainement dans ces représentations-là. Donc l'avantage de PISA c'est certainement d'avoir mis en lumière le fait que oui, il y a une corrélation entre performance et milieu d'origine. »⁶⁴⁷

La portée de l'outil statistique international donné par PISA permet d'objectiver, par une méthodologie complexe et unanimement reconnue, ce qui aurait pu être perçu auparavant pour de l'observation subjective, sans étalon de mesure de l'ampleur du problème. Deux paramètres sont par ailleurs nouveaux et contribuent à « révéler » sans cesse l'importance du problème public des inégalités scolaires en France, l'un étant propre à la fois à PISA et à CEDRE et l'autre étant propre à PISA. Le premier est la représentation chiffrée du renforcement de ces inégalités de cycle en cycle. Le second, et c'est sans doute celui dont l'importance surpasse les autres, est la mise en lumière de la place de plus en plus décevante de la France dans le palmarès des pays participants à PISA les plus inégalitaires. PISA ne révèle pas les inégalités scolaires à l'œuvre dans le système éducatif français, il montre leur aggravation et les révèle aux yeux du monde. Le changement de perspective, que l'on peut voir dans la loi portée par Vincent Peillon, était bien souligné en entretien par Agathe Cagé, ancienne conseillère au cabinet du ministre :

« C'est difficile parce que comme toutes les évaluations scolaires, ça mesure l'impact sur le niveau d'élèves qui ont 9 ans d'apprentissage. Là ce sont des enfants entrés en 2007. La source des problèmes est ancienne. Ce que PISA

⁶⁴⁷ Entretien avec Isabelle Robin, Directrice du département de la recherche et du développement, de l'innovation et de l'expérimentation, Dgesco, 7 avril 2017.

montre, et c'est ça qui est plus intéressant, c'est que la France fait beaucoup moins bien que les autres. C'est un système beaucoup plus producteur d'inégalités. »⁶⁴⁸

Par ailleurs, les implications des constats faits à partir des indicateurs de l'OCDE n'ont pas le même objet que celles faites à partir des travaux universitaires ou par les évaluations nationales menées par la DEPP. PISA incite à travailler non sur la réduction des inégalités en soi pour ce que cela dit du modèle de société dans lequel elles existent, mais plutôt parce que l'existence d'inégalités scolaires fortes réduit la performance d'ensemble du système éducatif. Le sujet de l'indicateur n'est pas l'élève, mais bien le système éducatif et ce que cela laisse préfigurer pour le fonctionnement à venir du marché du travail et de l'économie. C'est cette image dégradée auprès des partenaires internationaux qui donne une nouvelle dimension d'urgence au problème public des inégalités scolaires, dont le gouvernement acte la prégnance en reconnaissant l'inscription des enjeux nationaux de politique éducative dans un cadre élargi et international. Réciproquement, PISA a bénéficié largement de son parti-pris de concentrer ses efforts de communication autour de ses indicateurs sur les inégalités scolaires. En France, l'arrivée d'un gouvernement de gauche un an après la publication des résultats de PISA 2009 en a amplifié la voix.

Deuxièmement, le contexte politique et idéologique devient en 2012 très favorable à l'utilisation des indicateurs de PISA portant sur les inégalités scolaires. La gauche socialiste portée par François Hollande arrive non seulement à la Présidence de la République en mai 2012, mais elle est dans le même temps majoritaire aux deux assemblées, Assemblée nationale et Sénat. Ce triple alignement des forces politiques, à la fois dans exécutive et législative, construit un terrain on ne peut plus favorable au lancement d'une réforme d'ampleur du système éducatif autour d'un discours commun du constat du niveau dramatique des inégalités scolaires montré aux yeux de tous par PISA. Le soutien des parlementaires de gauche a été moteur de la montée en expertise de PISA, comme le soulignait Marie-Christine Blandin, présidente EELV de

⁶⁴⁸ Entretien avec Agathe Cagé, Conseillère au cabinet du ministre de l'Education Nationale en charge du second degré, des programmes et de l'évaluation des élèves, octobre 2016.

la commission des affaires culturelles et de l'éducation du Sénat au moment du vote de la loi.

« Quand on a fait la loi Peillon des années plus tard, dès le départ Peillon qui est un républicain quand même traditionnel entre guillemets mais avec tous les bons côtés de cette République, lui aussi a dit « mais il ne faut pas abandonner le socle, c'est un pacte République/parents/élèves, finalement faut garder le socle. Mais faut qu'il soit plus intégré à la distribution de la progression, de la discipline, des programmes, donc c'est pas quelque chose parallèle, c'est consubstantiel de l'école de base, c'est-à-dire depuis le début de la scolarité jusqu'à 16 ans ». Et cette loi Peillon s'est appuyée sur la contestation du bilan de la droite au travers des résultats de PISA dont l'énoncé accablant n'était pas le très bas niveau de la France, c'était le niveau moyen de la France, et surtout c'était l'école la plus inégalitaire de toutes. Et donc Peillon en a fait quelque chose de structurant pour sa loi, une école inclusive, une école qui n'accentue pas, ou plutôt qui corrige les inégalités. Et nous, comme je présidais la commission de la culture à ce moment-là, avec Françoise Cartron qui était rapporteure, on en a vraiment fait un axe à amplifier.

Donc c'était vraiment pour corriger ce mauvais bilan PISA, pas sur le niveau mais sur les inégalités. »⁶⁴⁹

Le raisonnement adopté par le gouvernement et les parlementaires de la majorité a finalement peu dérogé à la règle de construire une réforme à partir d'une critique classique du bilan des gouvernements précédents. La critique a porté ici principalement sur l'énoncé « accablant », selon les mots de Françoise Cartron, du niveau d'inégalités scolaires en France en soi, mais surtout par rapport aux autres pays participants. En accord avec les parlementaires de la majorité, la stratégie politique a consisté à centrer le discours sur ce bilan désastreux au regard de l'objectif premier de l'école républicaine, assurer la réussite de tous les élèves, plutôt que sur une critique du concept de socle commun introduit par la réforme de 2005. Cela a d'autant plus de sens que le socle commun a été, on l'a vu, largement rapproché des concepts popularisés par l'OCDE à la fin des années 1990 et matérialisés par l'évaluation PISA. Le ministre Peillon pouvait alors difficilement asseoir sa réforme du système éducatif

⁶⁴⁹ Entretien avec Marie-Christine Blandin, Sénatrice écologiste du Nord, Présidente de la Commission éducation et affaires culturelles du Sénat, 3 novembre 2016.

fondé sur le constat de l'inexactitude d'un raisonnement fondé sur la mise en place d'un socle de compétences.

Troisièmement, les solutions pratiques proposées par la réforme doivent être budgétairement acceptables et dans le champ des possibles des administrations en place⁶⁵⁰. Les solutions proposées ici par le ministre ne nécessitent pas de changements brutaux des lignes de travail de l'administration centrale, ni une hausse jugée déraisonnable de la dépense publique – en tout cas pas réellement différente des implications budgétaires des décisions traditionnellement prises par les gouvernements de gauche sur la hausse du budget de l'éducation nationale. Le seul élément perturbateur sera la refonte des rythmes scolaires dans l'enseignement du premier degré, mais bien que mentionné dans le rapport annexé au projet de loi il sera l'objet d'un décret séparé.

L'influence d'expertise allouée à PISA par la conjonction de l'ensemble de ces éléments de contexte se fait au détriment de l'expertise de la DEPP qui encore une fois pâtit de son cadre de travail national, décalé par rapport aux enjeux perçus par les décideurs politiques.

« PISA ça a ouvert les yeux au politique. Plus d'écart entre les élèves les moins favorisés et les plus favorisés, c'est vrai qu'il y a un problème. Il en va de l'image de la France auprès de l'OCDE et de la communauté internationale. On doit réagir, mettre en place des dispositifs. La France a beaucoup pris en compte PISA au détriment d'autres études, nationales, qui ont été menées par la DEPP. »⁶⁵¹

« C'est vrai que politiquement ça correspondait à un agenda parce qu'il y a eu un changement de majorité à ce moment là. La nouvelle majorité avait mis une priorité sur l'éducation quand même. Il y avait eu la grande concertation pendant l'été 2012, il y avait eu beaucoup de réflexion et donc on attendait les résultats PISA comme étant le bilan de ce que la majorité précédente avait pu faire. Et là, la question des inégalités, en France, avec une mesure du déterminisme social telle que PISA peut la faire, qui montrait que la France était le pays dans lequel les inégalités étaient les plus fortes, a

⁶⁵⁰ Erik Neveu, « La construction sociale des problèmes publics », *Idées économiques et sociales*, n°190, décembre 2017.

⁶⁵¹ Entretien avec Ronan Vourc'h, Adjoint au chef de bureau de l'évaluation des élèves, DEPP, 22 juin 2020. , souligné par nous.

déclenché disons une sorte de choc chez les politiques, qui a pu descendre jusqu'aux enseignants par l'intermédiaire des inspecteurs, par l'intermédiaire des formateurs, et notamment en mathématiques puisque c'était le domaine majeur à cette époque-là. »⁶⁵²

La DEPP avait pourtant notamment travaillé dès les premiers cycles de CEDRE sur la production d'indicateurs portant sur les inégalités scolaires. Les données fournies, si elles n'ont pas été du tout utilisées dans le cadre du débat sur la refondation de l'école, permettaient pourtant de regarder dans le temps l'évolution des inégalités scolaires selon plusieurs paramètres, comme les typologies de collèges (par exemple les collèges en secteur public de l'éducation prioritaire versus ceux hors secteur d'éducation prioritaire) laissant des possibilités explicatives diverses (comme l'effet de la mixité sociale dans les établissements, ou encore l'effet de la représentation des capacités des élèves dans les établissements⁶⁵³). Les regrets de ce non-usage des travaux de la DEPP, spécifiquement lors de la réforme de 2013, seront exprimés publiquement en 2016 à travers la publication d'un article dans la revue de la direction *Education & Formations*, écrit par Florence Lefresne de la mission aux relations européennes et internationales. Elle y affirme la force des enjeux qui entourent les évaluations des résultats des élèves qui contribuent à améliorer la connaissance et à permettre l'analyse du fonctionnement du système éducatif et de ses résultats, mais aussi à assurer une place dans la gouvernance européenne des objectifs internationaux assignés à l'éducation. Elle cite l'influence de PISA sur les décisions de réorientation du fonctionnement du système éducatif en France, mais rappelle que les conclusions données par l'OCDE avaient été également formulées dès 2011 par la DEPP, travaux à l'appui.

⁶⁵² Entretien avec Franck Salles, Chargé d'étude sur l'évaluation des élèves, National Project Manager, DEPP, 11 juillet 2017.

⁶⁵³ Voir par exemple les travaux d'Elodie Bellarbre, *Compenser les inégalités dans le secondaire en France : études à partir des enquêtes PISA, CEDRE et TALIS*, thèse pour le doctorat d'économie sous la direction de Jean-François Giret, soutenue à l'Université de Bourgogne Franche Comté le 19 octobre 2017.

« Les outils d'information et de mesure permettant ces comparaisons sont au centre d'enjeux importants: enjeux de connaissance (situer notre système éducatif au regard d'un ensemble de critères ; éclairer de façon fine ses points forts ou ses faiblesses) ; enjeux d'analyse (mieux comprendre les déterminants de variables essentielles telles que la performance ou l'équité) ; et enfin enjeux de gouvernance européenne (définir et mettre en œuvre les critères de référence et les indicateurs de suivi européens qui sous-tendent les orientations communes de la stratégie Éducation et Formation 2020). Ces enjeux font des données comparatives un outil à part entière dans l'aide au pilotage des systèmes éducatifs nationaux et dans l'aide à la décision politique. Au plan national, chacun peut avoir à l'esprit le rôle joué par PISA – fortement médiatisé – dans la prise de conscience collective des inégalités de performances à l'école et notamment des inégalités sociales de réussite scolaire. Cette prise de conscience a pu constituer l'un des leviers de la loi de refondation de l'École et des réformes qu'elle a engendrées, d'autant plus que les résultats de PISA 2009, puis de PISA 2012, convergeaient largement avec ceux des enquêtes nationales réalisées par la direction de l'évaluation, de la prospective et de la performance (DEPP) [Daussin, Kespaik, Rocher, 2011 ; Arzoumanian et Dalibard, 2015]. »⁶⁵⁴

La légitimité accordée aux indicateurs de PISA par les décideurs politiques dans les moments clés de réforme, au-dessus de celle de toute autre évaluation y compris nationale, a plusieurs implications sur le rôle d'évaluateur attribué à la fois à l'OCDE et à la DEPP. L'usage de PISA et l'inscription de cet usage dans un discours destiné non seulement aux citoyens et décideurs politiques français, mais aussi aux partenaires internationaux de la France participant à l'évaluation, font d'abord de l'OCDE un interlocuteur tout à fait nouveau à convaincre du bienfondé des réformes menées en France et des bons résultats obtenus. Dans le même temps, la DEPP, qui fonde sa stratégie de légitimation sur un rapprochement méthodologique de PISA et la valorisation de la maîtrise par ses agents de toutes les sortes d'évaluation du système éducatif utilisées pour piloter l'action publique, devient l'interlocuteur clé du gouvernement pour comprendre PISA et en communiquer les résultats y compris jusqu'au corps enseignant.

⁶⁵⁴ Florence Lefresne, « Du bon usage des comparaisons internationales dans l'aide au pilotage des systèmes éducatifs », *Education & Formations*, DEPP, 2016.

3) Le temps long de la réforme : l'expertise au service de la construction de l'évaluation des résultats de la réforme

La définition des modalités d'évaluation des résultats des réformes engagées, en 2005 comme en 2013, est intégrée dès le départ dans les débats législatifs, preuve de l'importance qu'elle revêt pour les ministres réformateurs et de l'appropriation générale des principes du Nouveau Management Public, y compris lorsque les projets de loi sont soumis par des gouvernements de gauche. La mise en débat parlementaire du processus d'évaluation inhérents à l'obligation de résultats qui vient remplacer l'obligation de moyens contribue à la dynamique d'institutionnalisation de l'évaluation⁶⁵⁵, et donc du rôle des évaluateurs, en officialisant le caractère incontournable de leurs travaux⁶⁵⁶ dans la conception d'une réforme.

a) Par quelle expertise répondre à la nouvelle obligation de résultat de la réforme ?

Lors de la réforme de 2005, l'importance du chiffre pour justifier l'action réformatrice apparaît au fil des interventions de François Fillon qui donne à sa réforme des objectifs chiffrés. Le ministre donne lui-même avant même le vote de la loi au Parlement les indicateurs de résultats qui seront examinés et à quelle échéance. Interrogé par un journaliste de *Métro* le 6 décembre 2007 sur la « valeur » que le ministre « *attache aux objectifs chiffrés* », François Fillon répond en s'inscrivant sans détour dans la logique promue par le Nouveau Management Public d'obligation non pas de moyens mais de résultats : « *ces objectifs chiffrés sont liés à notre volonté de mettre en place des systèmes d'évaluation pour vérifier que les buts sont atteints.* »⁶⁵⁷ L'obligation de résultats qui va de pair avec la mise en place du socle commun transparait à travers les impératifs d'efficacité qui sont assignés à la réforme, pour amener à une éducation de

⁶⁵⁵ Jacques Lagroye, Bastien François, Frédéric Sawicki, *Sociologie politique*, Paris, Presses de Sciences-Po, Dalloz, 2002.

⁶⁵⁶ Clément Lacouette Fougère, Pierre Lascoumes, « L'évaluation : un marronnier de l'action gouvernementale ? », *Revue française d'administration publique*, vol.148, n°4, 2013.

⁶⁵⁷ Interview de M. François Fillon, ministre de l'éducation nationale, de l'enseignement supérieur et de la recherche, dans "Métro" du 6 décembre 2004.

meilleure qualité pour tous les enfants. François Fillon l'exprime ainsi dans le document reprenant les 14 propositions de sa réforme : le socle permet « *d'instaurer une obligation de résultat et de moderniser le fonctionnement du service public de l'éducation* »⁶⁵⁸. Cette obligation a pour corollaire l'obligation de mettre en place un système d'évaluation des résultats de la loi. Le Haut conseil de l'éducation le souligne d'ailleurs dans ses *Recommandations pour le socle commun*, publiées en mars 2006⁶⁵⁹ : « *L'école a une obligation de résultats, effectifs et vérifiables* ». Le décret du 11 juillet 2006 vient officialiser le lien irréductible fait entre la mise en place du socle et la construction d'un mécanisme d'évaluation, considérant « *indissociable* » les deux exigences, de contenu du socle d'une part et de l'évaluation à mener sur les résultats de la définition de ce contenu d'autre part.

Cette volonté d'inscrire le pilotage par les chiffres dans la logique de la réforme portée par François Fillon était déjà énoncée dès les premières lignes de l'exposé des motifs du texte.

« Ce projet ambitieux appelle un ensemble de mesures qui ne relèvent pas toutes du domaine de la loi. C'est pourquoi le projet de loi propose au Parlement d'approuver un rapport annexé présentant de manière synthétique la politique que le Gouvernement entend conduire au cours des prochaines années et fixant des objectifs quantitatifs qui permettront de mesurer les performances du système éducatif. »⁶⁶⁰

Les parlementaires sont invités à non seulement travailler les articles du projet de loi comme ils peuvent le faire pour tout autre texte, mais aussi à donner leur assentiment sur la façon dont l'évaluation des objectifs de la réforme sera menée a posteriori. Le Haut Conseil de l'éducation prévu par le texte doit présenter un bilan annuel des résultats du système éducatif, et le gouvernement doit fournir un rapport sur l'évaluation de la mise en œuvre de la loi. L'évaluation du dispositif est ici pensée et construite en même temps que la création du dispositif.

⁶⁵⁸ Ministère de l'Éducation nationale, Loi d'orientation sur l'École. 14 propositions de François Fillon, novembre 2004, p.7.

⁶⁵⁹ Haut conseil de l'éducation, *Recommandations pour le socle commun*, 23 mars 2006.

⁶⁶⁰ Projet de loi d'orientation pour l'avenir de l'école n°2025, enregistré à la Présidence de l'Assemblée nationale le 12 janvier 2005.

Le mode d'évaluation des effets de la réforme fait d'ailleurs débat, la DEPP et les autres directions du ministère ne faisant pas office d'évaluateur suffisamment indépendant aux yeux de certains parlementaires. André Chassaing, député communiste, et son groupe proposent ainsi lors de l'examen du texte de créer un observatoire, voire plusieurs observatoires décentralisés. Il fait directement le lien lors de la défense de son amendement entre les défauts du système éducatif et l'impossibilité pour le ministère de fournir en interne une évaluation indépendante. Il explique ainsi son raisonnement :

« On parle de faillite du système éducatif, on tient des propos à l'emporte-pièce sur certains choix pédagogiques, qui peuvent être justifiés, on affirme qu'à certains stades du cursus scolaire, les inégalités se creusent et génèrent l'échec scolaire. Mais malgré tous les constats qui sont faits sur les insuffisances du système, on laisse l'évaluation aux seuls hauts fonctionnaires du ministère de l'éducation nationale. »⁶⁶¹

Les réserves émises par l'opposition ne suffiront pas à retirer à la DEPP la mission officielle d'évaluation des résultats de la réforme. Les objectifs du socle commun sont officiellement évalués par la DEPP, par le biais d'évaluations nationales et de dispositifs expérimentaux comme le livret personnel de compétences mis en place en 2007 et qui doit comporter une attestation de maîtrise des connaissances et compétences du socle commun. La direction du ministère est également chargée de fournir, dans le cadre de la LOLF en prévision des lois de finance, la proportion d'élèves maîtrisant les compétences du socle commun en fin de troisième, au palier 3⁶⁶² – soit environ à l'âge qu'ont les élèves participants à PISA. Cette contribution à la LOLF, qui occasionne des débats réguliers sur les missions de l'école et les résultats obtenus, permet certes de valoriser à échéance régulière les travaux menés par la DEPP. Mais elle a pour effet collatéral de lui conférer un rôle encore une fois descriptif d'un état du système éducatif à un instant t, au détriment du potentiel analytique des données fournies dans une démarche, comme son nom l'indique pourtant, prospective.

⁶⁶¹ 1^{ère} séance du 17 février 2005.

⁶⁶² Le socle commun est validé en trois étapes, appelées « paliers ». Le premier est atteint en CE1, le deuxième en CM2, et le troisième en fin de collège en classe de 3^e.

La DEPP communique sur la nature de la mission qui lui est confiée dans le cadre de la LOLF, en écrivant que ses « *indicateurs répondent à une exigence forte : la comparabilité dans le temps. En effet, ils doivent permettre de mesurer l'évolution des acquis des élèves – en termes de compétences de base – et l'atteinte des objectifs fixés. Ils seront calculés chaque année à partir des résultats obtenus par un échantillon d'élèves à des tests standardisés* »⁶⁶³.

Les données de la DEPP seront utilisées dans les documents préparatoires de la LOLF pour nourrir les débats budgétaires annuels, sans donner lieu à des réflexions prospectives sur les évolutions de fond souhaitables du système éducatif au vu de l'atteinte des objectifs chiffrés dérivés de la réforme de 2005. L'évaluation des résultats de la réforme sera, elle, réellement effectuée par le Haut Conseil de l'Éducation, qui formulera ses recommandations sur l'évolution du fonctionnement du système éducatif au moment de faire son bilan des résultats de l'École. Ils seront encore ici fondés sur un constat posé par la DEPP, et analysés à la lumière des conclusions de l'OCDE et notamment de PISA. Le rapport du HCE de 2011 détaille ainsi, en traitant le sujet de l'échec scolaire et de l'acquisition du socle commun, problème central de la réforme menée par François Fillon :

« Étant donné que 40 % des élèves sont en échec ou en difficulté (42), le Haut Conseil de l'Éducation est conscient qu'atteindre l'objectif de la maîtrise du socle commun par tous constitue un défi pour notre système éducatif. Pourtant, plusieurs pays de l'OCDE, notamment le Canada et la Finlande, ont réussi à relever le défi et démontrent que l'ambition du socle commun n'est pas irréaliste. »

Note de bas de page 42 : Ce chiffre repose sur les évaluations-bilans CEDRE (Cycle des évaluations disciplinaires réalisées sur échantillon), réalisées par la DEPP du ministère depuis 2003, pour mesurer les acquis des élèves au regard des objectifs fixés par les programmes disciplinaires à la fin de l'école primaire et à la fin du collège. »⁶⁶⁴

L'évaluation CEDRE n'est pas ici employée à la pleine mesure de ses capacités analytiques. Elle donne le constat du problème à résoudre, mais les perspectives de

⁶⁶³ Thierry Rocher, Sylvie Fumel, Jean-François Chesné, *Méthodologie de l'évaluation des compétences de base en français et en mathématiques en fin d'école et en fin de collège*, Note d'information n°37, DEPP, décembre 2008.

⁶⁶⁴ Haut Conseil de l'Éducation, *Mise en œuvre du socle commun. Bilan des résultats de l'École*, 2011.

solution à y apporter sont encore une fois issues de PISA. Alors que CEDRE reprend les codes de l'évaluation PISA, il n'arrive pas à trouver une place première dans les outils d'analyse et de prospective des instances chargées de formuler des jugements d'experts.

La réforme de 2013 s'inscrit une dynamique similaire de construction du dispositif d'évaluation au moment de la définition des objectifs nouveaux du système éducatif, mais réduit fortement le rôle d'évaluation assigné à la DEPP. Au-delà de l'évaluation des résultats des élèves qui leur est toujours confiée, l'évaluation des résultats du système éducatif se construit sans réellement intégrer la direction du ministère comme un interlocuteur légitime. On l'a vu⁶⁶⁵, l'indépendance de la DEPP est mise en cause dans les débats parlementaires portant sur l'article 21 créant le CNESCO, relativisant fortement la légitimité dont ses travaux peuvent sembler jouir à la lecture des données utilisées pour justifier des orientations du projet de loi. Elle est nommément mise en cause comme devant voir la qualité de ses travaux étudiées par l'instance évaluatrice créée par la loi. Il est intéressant de voir que les missions du CNESCO concernant les évaluations internationales et en particulier PISA, indiqué comme l'évaluation centrale au niveau international, sont très différents et témoignent d'une crédibilité accordée tout à fait différente. Alors que dans le cas de la DEPP le CNESCO serait chargé de « *se prononcer sur les méthodologies et les instruments des évaluations conduites*⁶⁶⁶ », l'évaluation de l'OCDE ne se voit pas l'objet d'une défiance à l'égard de sa méthode. Il aurait été difficilement explicable, on l'imagine, de fonder sa réforme sur les données d'une étude dont on dirait, dans le même texte de loi, que la méthodologie ayant permis de les obtenir devrait être vérifiée par le CNESCO. Le CNESCO se voit bien missionné spécifiquement sur PISA. Son rôle est alors d'en prendre connaissance et de se prononcer sur les enjeux des résultats de PISA, et non sur le bienfondé de la construction des résultats et leur significativité dans cadre national français. Pour la commission sénatoriale rendant son rapport après audition des « *experts (sociologues, économistes, spécialistes de sciences de l'éducation) rencontrés par la*

⁶⁶⁵ Voir chapitre 5.

⁶⁶⁶Rapport n°568 fait au nom de la commission de la culture, de l'éducation et de la communication *sur le projet de loi d'orientation et de programmation pour la refondation de l'école de la République*, par Françoise Cartron, sénatrice, enregistré à la Présidence du Sénat le 14 mai 2013, p.110.

*rapporteuse*⁶⁶⁷ » (dont il faut rappeler que la DEPP ne fait pas partie), la direction du ministère serait en fait un « *opérateur* », et le CNESCO le vrai évaluateur. Le rapport de l'Assemblée nationale va même plus loin, en considérant que la DEPP est un évaluateur « non indépendant », et que les autres instances d'évaluation du système éducatif sont en fait soit externalisées comme l'OCDE, « *c'est-à-dire PISA*⁶⁶⁸ » et la Cour des comptes, soit mobilisées par des missions autres. Yves Durand, rapporteur socialiste⁶⁶⁹ du texte pour la commission, vise ici spécifiquement le HCE chargé de fournir un avis sur l'élaboration des programmes et la formation des enseignants, en plus de son regard sur les résultats du système éducatif. Le rapporteur valorise néanmoins CEDRE malgré son avis tranché sur la légitimité d'expert de la DEPP, en estimant que ses indicateurs sont peu connus mais rigoureux, et sont à même de remplacer avec pertinence les indicateurs de PISA. Ce positionnement paradoxal quant au caractère d'expertise de la DEPP amène à constater deux choses. D'une part, le débat sur l'identité et donc l'indépendance de l'expert évaluateur prend le pas sur celui des méthodes valides d'évaluation. Le questionnement autour de l'indépendance de la DEPP et la légitimité accordée dans le même temps à l'OCDE rend compte de la crainte supérieure des parlementaires de voir le ministère piloter en sous-main les évaluations nationales des résultats du système éducatif de la DEPP, alors même que la DEPP est la seule interlocutrice de l'OCDE participant l'élaboration de PISA pour la France. Ce travail ne semble pas poser question sur la validité des indicateurs produits par PISA dans le cas français. D'autre part, la reconnaissance de l'expertise produite par CEDRE, caractérisé par le rapporteur par sa méthode et son antériorité proches de celles de PISA mais son raisonnement disciplinaire plus adapté aux objectifs fixés en termes d'acquis des élèves, montre que la stratégie de la DEPP de crédibiliser ses travaux dans le sillage de PISA commence à porter ses fruits, mais

⁶⁶⁷ *Ibid.*, p.111.

⁶⁶⁸ Rapport n°767 fait au nom de la commission des affaires culturelles et de l'éducation *sur le projet de loi d'orientation et de programmation pour la refondation de l'école de la République, Tome 1*, par Yves Durand, député, enregistré à la Présidence de l'Assemblée nationale le 28 février 2013, p. 364.

⁶⁶⁹ Yves Durand est à l'époque affilié au parti socialiste, qu'il quittera en 2017 pour rejoindre La République en Marche, le parti fondé par Emmanuel Macron.

manque encore de popularité auprès des décideurs politiques de droite comme de gauche.

Le choix est par ailleurs fait de créer de nouvelles instances d'évaluation, et notamment au sein des assemblées parlementaires, qui viennent s'ajouter aux instances indépendantes prévues par les textes de loi. Yves Durand, rapporteur de la loi pour la refondation de l'école, les décrivait ainsi en entretien et précisait leur façon de mener leurs missions.

« Alors il y a le comité d'évaluation de l'Assemblée nationale, pour lequel j'ai fait un rapport. Ça, c'est le premier. Et la deuxième, c'est la présidence du comité de suivi de la loi de refondation de l'école qui n'est pas véritablement de l'évaluation. Comme son nom l'indique, c'est du suivi. Mais qui est néanmoins à la marge de l'évaluation. C'est vraiment un travail d'évaluation. On a six mois. Y'a un comité d'évaluation qui est déjà en phase d'ailleurs avec les administrateurs de l'assemblée, qui sont totalement institutionnels. Et on a aussi un budget, on a les moyens, notamment des moyens d'investigation, des moyens de déplacement. Dans ce cadre-là, on a été dans plusieurs académies, on a été dans plusieurs pays étrangers, on était en Belgique et aux Pays-Bas. Et on peut faire travailler des cabinets d'audit privés qui eux-mêmes, d'ailleurs c'est assez intéressant, ont leur propre regard. Par exemple, on avait pris un cabinet qui a travaillé sur trois académies autres que celles sur lesquelles nous travaillions.

L'OCDE intervient à plusieurs moments. On les a interrogés, on les a auditionnés. Ils nous ont donné un certain nombre de résultats de leurs travaux. On a pu croiser un certain nombre de résultats, voire de préconisations. »⁶⁷⁰

La DEPP est laissée à l'écart des travaux menés dans le cadre des activités des instances évaluatrices parlementaires, alors même que le jugement d'expert de l'OCDE est ouvertement pris en compte dans les conclusions des parlementaires, et qu'il est complété par ailleurs par des cessions de portions d'évaluation au secteur privé. Alors que par le passé la DEPP avait pu évaluer des dispositifs complétant le socle commun, y compris en travaillant aux côtés de cabinets privés comme dans le

⁶⁷⁰ Entretien avec Yves Durand, Sénateur socialiste puis LREM de Mayenne, rapporteur de la loi refondation de l'école, président du comité de suivi et membre du comité d'évaluation des politiques publiques, 14 juin 2017.

cas du livret de compétences expérimental⁶⁷¹, elle n'est cette fois pas sollicitée. Les comités d'évaluation et de suivi de l'Assemblée nationale ont fait appel, selon Yves Durand, à des cabinets d'audit qui ont travaillé sur les résultats de la réforme dans plusieurs académies. Et cela alors même que des évaluations spécifiques auprès d'un échantillon de 160 000 élèves ont été par exemple menées en 2015 par la DEPP⁶⁷², pour pouvoir évaluer les écarts de performance éventuels entre les académies. Leurs données produites pour alimenter les conclusions, notamment en termes d'égalité territoriale de mise en place du socle commun, ne sont ni popularisées par la communication ministérielle ni valorisés dans les débats parlementaires. A la fin de la période à laquelle nous nous intéressons, la DEPP se détache difficilement de son étiquette d'évaluateur du ministère, y compris avec son évaluation CEDRE dont la qualité et le potentiel sont reconnus par ailleurs par les parlementaires.

b) Le recours à l'expertise de PISA : l'apparition d'un nouvel évaluateur à convaincre

Le recours systématique à l'expertise de l'OCDE et de son PISA a pour effet concomitant de faire apparaître un nouvel évaluateur, un « *œil global* ⁶⁷³», dont les avis sont scrutés à la fois par les décideurs politiques français, majorité comme opposition, et par leurs homologues des autres pays participants. Les réformes entreprises le sont certes pour répondre à des constats jugés problématiques au niveau national, à un moment favorable pour tenter de le faire. Mais elles le sont aussi pour montrer au monde que l'on est capable de prendre conscience des problèmes publics grâce à PISA et de réagir une fois le problème posé. La réforme ne doit plus seulement être évaluée par les instances évaluatrices nationales, qu'elles pré-existent à la loi comme la DEPP

⁶⁷¹ La DEPP a travaillé sur l'évaluation du LCEX, prévu par la loi relative à l'orientation et à la formation tout au long de la vie de 2009, aux côtés du cabinet ASDO chargé du volet qualitatif de la démarche évaluative de l'expérimentation, menée dans 164 collèges et lycées sous la responsabilité à la fois des ministères de l'Éducation nationale et de l'Agriculture.

⁶⁷² DEPP, *Évaluation numérique des compétences du socle en début de sixième : des niveaux de performance contrastés selon les académies*, Note d'information n°18, 2016.

⁶⁷³ Antonio Novoa et Tali Yariv-mashal, « Comparative Research in Education : a mode of governance or a historical journey ? » *Comparative Education*, vol. 39, 4, 2003.

ou qu'elles soient créées par la loi. Elle devient objet d'évaluation pour l'OCDE, qui, rassembleuse des réactions de politique publique à ses recommandations, donne à voir la capacité réformatrice des participants à PISA et les effets des réformes sur les résultats des systèmes éducatifs.

La prise de conscience par les décideurs français du déplacement du regard évaluateur et de la reddition de compte déplacée vers l'OCDE est formulée assez rapidement après le lancement de PISA. Gérard Longuet le 15 mars 2005, lors de l'examen du texte de la réforme de 2005, mettait déjà en avant cette « observation de l'extérieur » des résultats de l'éducation en France engendrée par la participation à une évaluation internationale particulièrement regardée par la France comme par ses partenaires étrangers.

«La commission des finances souhaite faire une observation : nous ne nous situons plus dans un débat franco-français. Naturellement, d'une majorité à l'autre, on peut se rejeter les responsabilités. La vérité est que nous sommes observés de l'extérieur, notamment en raison de notre participation au programme international de suivi des acquis, puisque le ministère de l'éducation nationale transmet des données. Un tel programme permet d'établir, année après année, la performance du système scolaire et le savoir moyen des élèves français au regard de leurs jeunes homologues des pays de l'OCDE. »⁶⁷⁴

PISA s'inscrit dans une démarche de panoptisme global⁶⁷⁵ institutionnel, en fournissant aux pays participants à PISA les informations issues de l'observation des systèmes éducatifs de leurs partenaires-concurrents, dont ils peuvent tirer profit dans le cadre de la compétitivité qui s'exerce au niveau mondial en matière d'éducation et de formation. La démarche de surveillance⁶⁷⁶ de l'OCDE par PISA amène à ce qu'elle se voit attribuer le rôle de diffuseur des résultats de chaque pays participant auprès des autres, par le biais de ses palmarès. Ce phénomène nouveau rend presque obsolètes les techniques classiques des débats d'opposition-majorité se rejetant mutuellement la

⁶⁷⁴ Intervention en séance de Gérard Longuet, sénateur UMP de la Meuse, le 15 mars 2005.

⁶⁷⁵ Bob Lingard, « Le « panoptisme global » : l'influence des tests internationaux et des infrastructures de données en éducation », *Administration & Éducation*, vol. 145, n°1, 2015.

⁶⁷⁶ Bob Lingard, *art. cit.*, Bob Lingard, Wayne Martino et Goli Rezai-Rashti, « Testing Regimes, Accountabilities and Education Policy : Commensurate Global and National Developments », *Journal of Education Policy*, vol.28, 5, 2013.

faute à la publication des évaluations des résultats du système éducatif, et oblige à fonder les réformes sur plus qu'un rejet des politiques antérieures – même si on l'a vu PISA fournit tout de même suffisamment d'indicateurs pour que chaque tendance politique y retrouve ses marqueurs idéologiques et puisse les utiliser pour justifier ses prises de position. La persistance du constat de dégradation des résultats de l'éducation en France donnée par PISA, et cela malgré les alternances politiques qu'a connues la France depuis le début des années 2000, aboutit par ailleurs mécaniquement à une dévalorisation de la puissance argumentative du rejet de la faute sur le gouvernement précédent.

Les données de PISA sont suivies de recommandations de l'OCDE, qui se place, on l'a vu, dans une posture de conseil, notamment aux moments clés de réforme. Le fondement même de la loi sur un constat formulé par l'OCDE dans les publications PISA engendre *per se* la nécessité d'évaluer l'efficacité de la réforme par les indicateurs de PISA. L'évaluation classique du système éducatif, que cela soit à partir des résultats des élèves qui y sont, qui y sont passés, ou à partir d'indicateurs à retardement comme l'entrée des jeunes actifs sur le marché du travail et l'adéquation de leurs compétences avec les emplois à pourvoir, ne permettent plus de déterminer avec des chiffres comparables si les résultats sont meilleurs ou moins bons qu'au moment de la réforme. L'œil global ne s'arrête pas à l'observation des résultats de l'éducation mais regarde aussi a posteriori le type et la pertinence des actions menées par les gouvernements, en plus d'en évaluer les résultats au moyen des cycles suivants de PISA. Son expertise dépasse la fonction d'aide à la décision, et est actée par les décideurs publics qui en font, comme Vincent Peillon, la base de leur réforme. Celui-ci

« M. Vincent Peillon, ministre de l'éducation nationale. Vous avez raison, monsieur le député, de me demander si la priorité accordée à l'école depuis dix-huit mois suffit. J'ai rappelé au président Borloo, voici un instant, la priorité donnée à l'école primaire, la remise en place de la formation des enseignants, les réseaux consacrés aux « décrocheurs », la mutation de notre système d'orientation, la réforme du temps scolaire saluée ce matin par l'OCDE »⁶⁷⁷

⁶⁷⁷ Réponse du ministre Vincent Peillon à la question posée au gouvernement n°1402 par Vincent Feltesse, socialiste, publiée au journal officiel de l'Assemblée nationale le 14 décembre 2013.

L'OCDE s'approprie pleinement ce nouveau rôle, utilisant les résultats PISA pour les publications de la série « *Better policies* », lancées au début des années 2010, qui jugent en toutes lettres le bienfondé ou non des réformes mises en place en se focalisant sur certains pays, dont la France. L'OCDE met en place avec cette publication une vraie logique d'évaluation des réformes menées dans les pays, sortant du rôle d'évaluateur des résultats des élèves. La publication reprend largement les recommandations sous-jacentes aux conclusions des cycles PISA, renforçant encore le phénomène de sur-référence rendant PISA inévitable dans l'ensemble des champs de politique publique. L'édition 2015 du rapport évalue ainsi très directement la réforme menée par Vincent Peillon.

« Dans ce contexte, la Loi d'orientation et de programmation pour la refondation de l'école de la République du 8 juillet 2013 visant à lutter activement contre le décrochage et l'échec scolaire est la bienvenue. La mise en œuvre récente de nombreuses réformes qui en découlent dans le primaire et au collège pourrait ainsi répondre, en fonction des modalités de leur mise en œuvre, à certains enjeux actuels et participer à améliorer les résultats et la formation des élèves. D'autant plus que des réformes mises en place par d'autres pays de l'OCDE, se fixant elles aussi comme priorité la réduction des inégalités sociales, n'ont en aucun cas conduit à un nivellement par le bas des performances. Bien au contraire, la proportion d'élèves en échec scolaire a en général reculé en moins de 10 ans dans ces pays, alors même que celle des bons élèves a augmenté. Les réformes actuelles vont dans la bonne direction et doivent être approfondies. »⁶⁷⁸

Face à cette nouvelle instance d'évaluation fondée sur des données statistiques, l'enjeu de répondre à PISA devient clé et c'est dans cet interstice d'influence que le rôle de la DEPP trouve l'une de ses missions centrales aux yeux des décideurs politiques, pour au moins deux raisons. D'abord parce que répondre à PISA signifie d'abord suffisamment répondre à PISA. Les bons taux de participation des établissements sont une des conditions indispensables pour que les résultats de la

⁶⁷⁸ OCDE, *Politiques meilleures, France*, Editions OCDE, 2015, p. 2.

France soient significatifs au moment où ils seront traités par le consortium puis présentés sous la forme d'indicateurs par l'équipe PISA-OCDE. La DEPP, en tant qu'opératrice nationale des tests, a pour mission d'assurer un taux de réponse maximal, et doit pour cela communiquer largement sur les enjeux de l'évaluation auprès des chefs d'établissements tirés dans l'échantillon des participants au test. La communication insiste, à l'échelle des établissements, sur le besoin d'obtenir de bons taux de réponse. A l'échelle des élèves, elle doit inciter à comprendre l'enjeu de la participation à PISA et donc à réaliser les tests de la façon la plus sérieuse possible. L'arrivée de Jean-Marc Blanquer à la tête du ministère en 2017 marque le lancement d'un renforcement important de la communication vers les établissements participants, sollicitant pleinement les liens de la DEPP avec le monde enseignant dont ses agents sont en partie issus.

« Le ministre a voulu que PISA soit beaucoup plus porté en France que ça ne l'était. Nous, on envoyait des cahiers d'évaluation dans les établissements et voilà. On relançait quand même pour avoir des taux de réponse suffisants parce que sinon l'OCDE pouvait nous exclure. Mais quand même quand on voyait les réponses qu'on avait on s'interrogeait sur l'investissement des élèves et des établissements. Le ministère nous a demandé de faire une énorme com' sur PISA. Du coup on a fait une campagne nouvelle, le ministre est même venu. Il a expliqué aux chefs d'établissement qu'on allait faire ça bien, que c'est important pour la France, c'est l'étendard, c'est une référence. »⁶⁷⁹

L'insistance du ministre sur l'importance des résultats à PISA auprès d'élèves dont la scolarité n'est pas impactée par leurs réponses au test est fondamentale au vu de l'effet sur la performance des élèves de la perception de l'enjeu de bien répondre. La DEPP est sollicitée pour travailler sur le renforcement de la motivation des élèves de l'échantillon PISA et publie même un article sur le sujet en 2017, dans la revue *Education & Formations*, intitulé « La motivation des élèves à répondre à un test

⁶⁷⁹ Entretien avec Ronan Vourc'h, Adjoint au chef de bureau de l'évaluation des élèves, DEPP, 22 juin 2020.

standardisé. »⁶⁸⁰ Elle montre que dans le cadre d'évaluations à faible enjeu pour les élèves qui y participent, comme c'est le cas pour CEDRE ou PISA, l'implication et la motivation des élèves tend à être moins importante que lorsque les évaluations ont un enjeu perçu comme élevé.

La DEPP est par ailleurs l'interlocutrice privilégiée de l'OCDE, et elle permet en cela de faire le pont entre la communauté enseignante d'une part et l'organisation internationale d'autre part. Ce lien aux deux bouts de la chaîne, OCDE d'un côté et enseignants de l'autre, valorisé en interne à la direction, est utilisé par les cabinets en place pour travailler à la compréhension et à la bonne interprétation médiatique des résultats de PISA. La DEPP, à défaut de produire une analyse approfondie des indicateurs PISA qui est déjà largement réalisée par l'OCDE dans ses publications, se charge de mieux organiser la communication des résultats fournis par l'OCDE auprès des médias et des enseignants. Les nouvelles formules de conférence de presse préalables à la publication des résultats, ayant pour but de permettre aux journalistes de mieux saisir le fond des données PISA, dont on a vu qu'elles permettaient à la DEPP de reprendre la main sur une partie de la communication cadrée largement par ailleurs par l'OCDE, rentrent dans les missions clés de la DEPP à partir du cycle PISA 2012. La direction s'évertue également à porter les résultats PISA jusqu'aux enseignants et au plus près de la pratique d'enseignement, en intervenant notamment dès 2013 sur les temps de formation des professeurs au sein des ESPE.

Ces missions construites dans le cadre de la gestion par la DEPP de PISA permettent à la direction de se positionner en référente de l'utilisation des données de l'évaluation internationale et de ce qu'elles permettent d'imaginer en termes de politique éducative. C'est particulièrement le cas avec l'arrivée de Catherine Moisan à la tête de la DEPP en 2013, qui tient à maîtriser la communication autour de PISA et à en diffuser les résultats auprès à la fois du ministre et des directions opérationnelles de l'administration du ministère.

⁶⁸⁰ Sylvie Fumel, Saskia Keskaik, « La motivation des élèves à répondre à un test standardisé : Résultats d'une étude dans le cadre de Cedre compétences langagières et littératie. *Éducation & formations*, DEPP, 2017.

« C'est difficile parce que la publication des résultats appartient à l'OCDE. Du coup on est tenus au confidentiel absolu jusqu'au moment où ça sort. Nous on travaille sur les données, mais pour les ministres c'est pas facile de faire qu'ils ne causent pas. J'avais les résultats sur le déterminisme social donc j'avais demandé des réunions avec Vincent Peillon. Je savais ce qui allait se passer, la France déclinait sérieusement. Et ça a commencé à bouger. J'ai eu beaucoup plus de mal avec la Dgesco, à convaincre la direction qu'il fallait bouger. Ça a quand même déclenché des trucs sur les maths, mais ça n'a pas eu l'effet que ça a eu en Allemagne en 2001. »⁶⁸¹

Sa façon d'envisager le rôle de gardien et de passeur de l'information donnée par PISA est structurante dans la façon dont les missions de la DEPP ont évolué au cours des années 2010, et dont elles seront pérennisées même après son départ de la direction en 2015. Ancienne enseignante appartenant au corps des inspecteurs généraux, siégeant au conseil d'administration de l'ESPE de Créteil, ayant par le passé travaillé au sein de plusieurs cabinets ministériels et ayant notamment contribué largement à la mise en place des ZEP, Catherine Moisan permet, par son parcours, de populariser PISA aussi bien auprès du politique et de la haute-administration publique que du monde enseignant. Elle mène son travail d'explication de PISA avec Thierry Rocher, qui sera systématiquement associé à ses interventions, notamment auprès des formateurs en ESPE.

Ce positionnement auprès de tous les acteurs impliqués directement ou indirectement dans PISA, redonne à la DEPP l'une de ses missions centrales de relai légitime des données PISA auprès de ses interlocuteurs. Elle devient experte de l'explication des indicateurs PISA, experte de leur transmission aux acteurs de terrain, sans toutefois parvenir à se hisser au rang des experts privilégiés apportant publiquement leurs propres savoirs statistiques et les recommandations afférentes pour construire les nouvelles orientations de la politique éducative française.

⁶⁸¹ Entretien avec Catherine Moisan, Directrice de la DEPP de 2013 à 2015, 10 juin 2020.

Conclusion du chapitre

L'étude des deux grands moments réformateurs de la période 2000-2017 nous permet de compléter l'analyse des résultats de la stratégie initiée par la DEPP d'asseoir son rôle d'évaluateur clé aux moments de réforme. Ses possibilités d'intervention sont triples: d'abord dans la justification du bienfondé de la réforme grâce à ses travaux et aux problématiques d'action publique qu'elle cherche à soulever, ensuite dans l'accompagnement et la rationalisation des propositions d'amélioration de la réforme émises par les parlementaires dans le cadre des débats en séance sur les projets de loi, enfin dans la construction des outils d'évaluation des résultats de la réforme en termes de résultats des élèves.

On perçoit d'abord, à travers les deux moments de construction du fondement des lois de 2005 et de 2013, l'évolution des modes d'influence respectifs des travaux de la DEPP et de l'OCDE et de PISA pour orienter les problèmes publics à traiter par la réforme. La DEPP se trouve toujours en toile de fond de la construction des problèmes publics abordés par les réformes sans pour autant arriver à asseoir son rôle d'expert de l'évaluation de la politique éducative capable de fournir les données adéquates pour justifier l'urgence de traiter certains problèmes publics jugés prioritaires par les gouvernements réformateurs, mais aussi pour émettre un jugement d'expert entendu et écouté pour construire les solutions à y apporter. Valorisée indirectement en 2005 par la nomination de son ancien directeur Claude Thélot à la tête de la commission du débat national et la reprise de certains de ses résultats pour en alimenter les travaux, elle ne parvient pas à installer durablement sa légitimité d'instance évaluatrice capable d'apporter une expertise indépendante au moment de la fabrication de la loi et des nouvelles politiques publiques qu'elle contient.

A côté de cela, par le regard externe qu'elle permet et la reconnaissance large de la qualité méthodologique de ses travaux, l'OCDE est finalement utilisée avantageusement par les ministres réformateurs pour déplacer les termes du débat à un niveau international susceptible d'apparaître comme plus objectif que les données internes au ministère. Ce changement de positionnement de l'œil évaluateur des résultats de la politique éducative française permet non seulement de réduire le risque

de voir émerger des oppositions à la réforme de la part des acteurs de la communauté éducative et/ou de leurs représentants syndicaux, mais aussi de montrer aux partenaires internationaux de la France, dont la majorité participent à PISA, l'importance accordée par le gouvernement au niveau et à la qualité de la formation initiale des jeunes Français et donc à l'avenir de la compétitivité du pays. En 2013 en particulier, malgré les travaux nombreux menés par la DEPP sur le sujet des inégalités scolaires au centre de la réforme voulue par François Hollande et Vincent Peillon, la capacité offerte par PISA de remporter l'adhésion de la droite à la réforme dépasse tout autre usage des évaluations nationales, aussi intéressantes et adaptées à la façon française d'évaluer les résultats de l'éducation soient elles.

La sollicitation permanente et ouverte des données de PISA dans toutes les phases de la réforme renforce ce qui avait été initié par la réforme de 2005 fondée sur le besoin d'intégrer dans la politique éducative française les enjeux formulés par les stratégies européennes soutenues par la France : l'institutionnalisation de l'évaluation non seulement des résultats mais aussi des moyens entrepris pour les atteindre par l'OCDE. Si les résultats observés pour évaluer l'effet des réformes menées dans le champ éducatif sont toujours ceux des acquis des élèves via PISA, les moyens sont eux regardés aussi bien dans le champ de la politique éducative que dans celle des politiques de la formation professionnelle, sociale ou encore du travail et de l'emploi.

La DEPP ne parvient pas à se hisser au rang des experts sollicités prioritairement et publiquement par les décideurs politiques ni pour construire les nouvelles orientations de politique éducative ni pour en être l'évaluateur clé. L'usage répété de PISA, par l'ensemble des groupes parlementaires et par les gouvernements de toutes couleurs politiques, la contraint à se saisir de l'évaluation internationale pour valoriser ses propres travaux. En cela, elle réussit néanmoins à asseoir son rôle d'experte technique des évaluations nationales et internationales. Sa vision d'ensemble sur les résultats des évaluations menées par l'OCDE, les autres organisations internationales comme l'IEA et par elle-même, associée à des nominations au poste de directeur de personnalités en lien étroit avec le monde de l'enseignement à partir des années 2010, lui permet de faire circuler les conclusions de PISA complétées par celles de ses propres travaux jusqu'aux enseignants. Alors qu'ils n'étaient pas

considérés comme les destinataires de l'évaluation internationale PISA donnant à voir les résultats de l'action publique en éducation et non les résultats des pratiques d'enseignement, ils deviennent des interlocuteurs récurrents de la DEPP qui porte jusqu'à eux, à travers notamment les établissements de formation des enseignants et les colloques organisés par les acteurs du champ scolaire, les recommandations formulées par PISA et confirmées ou relativisées par ses travaux. Catherine Moisan, à la tête de la DEPP pendant la réforme Peillon et sa mise en œuvre, contribuera largement à accélérer ce mouvement. Fabienne Rosenwald, qui prend la direction de la DEPP en 2015, continuera d'ailleurs à entretenir et à renforcer, avec Thierry Rocher garant de la continuité du travail de la direction, le lien tissé avec le monde enseignant. Elle soutient à plusieurs reprises la nécessité de faire parvenir les résultats de PISA jusqu'aux enseignants, et fait de PISA non pas un outil de pilotage politique du système éducatif mais bien un outil guide pour les enseignants dans leur pratique pédagogique. Citée par le Conseil Scientifique de l'éducation nationale en 2020, elle estimera ainsi que « *les évaluations standardisées des élèves ont pour objectif de fournir aux enseignants des outils afin d'enrichir la pratique. Elles n'ont du sens que si elles sont utiles aux enseignants* »⁶⁸². Elle ouvre, avec sa déclaration, la possibilité pour la DEPP de devenir l'acteur charnière de la mise en politique publique des résultats de PISA par le vecteur du travail des pratiques d'enseignement et des conditions d'apprentissage des élèves, sans nécessairement impliquer de passer par de grands moments législatifs.

⁶⁸² Conseil scientifique de l'éducation nationale, *Présentation et activités 2019-2020*, Ministère de l'éducation nationale, 2020, p. 16.

Conclusion générale

Cette thèse porte sur l'émergence d'une forme d'évaluation des politiques éducatives, l'évaluation PISA menée par l'OCDE, et les effets de l'irruption d'un outil d'un genre nouveau dans l'organisation des acteurs du champ des politiques publiques éducatives en France et en particulier ceux chargés des évaluations statistiques nationales pour le ministère de l'Éducation nationale.

PISA est mis en place au milieu des années 1990, à un moment où l'évaluation des résultats de la politique éducative est encore peu institutionnalisée en France. La DEPP a été tout juste créée en 1987, dans le mouvement de développement d'un « souci de soi » de l'Etat incitant l'administration à construire des outils d'évaluation de l'action publique. Elle est interne au ministère de l'Éducation nationale, et a en charge de mener les évaluations nationales des résultats des élèves et de donner à voir toute une série d'indicateurs descriptifs sur le fonctionnement et l'organisation de l'enseignement en France. Ses méthodes de travail sont exclusivement statistiques, et s'inscrivent dans le courant de l'« evidence-based policy », conférant à la méthode statistique le pouvoir de représenter le réel avec la plus grande justesse. Le choix de la forme d'évaluation menée et des indicateurs produits témoigne des représentations que véhiculent les décideurs politiques de ce que doit être l'action de l'État. La discussion sur la validité technique des indicateurs de double d'un questionnement sur les implications que la construction de tel ou tel indicateur a sur l'action que doit engager l'État. Lorsqu'un indicateur est par exemple publié sur le lien entre origine socio-économique et résultats scolaires, comme cela est le cas dans PISA et certains travaux de la DEPP, il est non seulement implicitement signifié que cela est du ressort de l'intervention publique mais il intime également une réaction adaptée aux résultats obtenus.

C'est dans ce contexte que le projet PISA est proposé par l'OCDE aux pays membres. Alors que l'organisation n'a pas de prérogative particulière en matière de politique éducative puisque le contexte de sa création après-guerre l'a positionnée comme référence internationale des politiques économiques et financières, elle imagine une évaluation des résultats de l'éducation d'un genre nouveau. Il s'agit de donner à voir aux pays la façon dont les élèves en âge de quitter l'école obligatoire sont dotés des compétences indispensables pour mener une vie d'adulte, de citoyen et d'actifs accomplis. A revers des évaluations nationales telles qu'elles sont alors menées par la DEPP, l'OCDE s'affranchit des programmes scolaires et des niveaux d'enseignement pour monter un projet tourné vers une définition nouvelle de l'école, pourvoyeuse de compétences et non de connaissances.

L'enjeu de la définition du cadre d'analyse et des items présentés aux élèves participants à l'évaluation est crucial. L'OCDE crée les conditions d'un travail par consensus des pays volontaires pour participer à PISA, en formalisant différents groupes de travail internationaux dont elle est la pilote. Des groupes experts techniques chargés de définir les contenus des tests et les règles pour évaluer les réponses des élèves, de doublent de groupes de pilotage constitués de représentants des différents pays, et un référent est nommé dans chaque ministère de l'Éducation. En France, c'est la DEPP qui est choisie pour participer à PISA et représenter les intérêts de la France. Si les raisons de la participation à PISA n'ont pas été réellement mises à jour par l'étude documentaire effectuée dans cette thèse, plusieurs témoignages d'agents de la DEPP et de représentants du ministère en poste au moment du lancement de l'évaluation ont relevé le caractère perçu comme obligatoire de la participation demandée à la direction. Plusieurs causes conjointes ont été évoquées. La France est à l'époque très investie dans la construction d'un espace européen de l'éducation et de la formation. La mise en place d'une évaluation internationale de grande envergure, proposée par une organisation internationale dont les sujets de travail sont principalement comme son nom l'indique le développement économique et la coopération dans ce but, est alors aux yeux des décideurs politiques vue comme

une pierre supplémentaire à l'édifice. La participation à PISA relève aussi des enjeux de politique étrangère et diplomatique, en témoignera d'ailleurs les contacts pris par le ministère des Affaires étrangères avec la DEPP lorsque les résultats seront jugés trop mauvais et dévalorisant l'image de la France. La localisation du siège de l'OCDE à Paris est aussi évoquée comme l'une des raisons ayant fortement incité la France à participer, encore une fois pour des raisons diplomatiques et d'image internationale. Le pays avait déjà eu une mauvaise expérience avec une évaluation menée par l'OCDE dans le champ de la compétence en lecture des adultes (IALS), et il était alors perçu comme diplomatique compliqué de refuser deux participations de suite. L'engagement de l'OCDE à construire une évaluation avec les experts internationaux de tous les pays volontaires a par ailleurs contribué à atténuer les craintes d'un « deuxième IALS », sans toutefois les faire disparaître. Quelles que soient les causes de cet engagement, pour les agents de la DEPP, le discours est clair : la DEPP n'a pas pu faire autrement que de participer, malgré des alertes portées auprès du gouvernement sur le risque de non-adaptation du modèle de l'OCDE au système français.

L'OCDE se constitue dès lors une « arène de la gouvernance éducationnelle » où elle est à la fois l'arbitre des débats nationaux et le décideur en dernière instance. Elle oriente les choix en proposant des axes de travail, parmi lesquels les pays peuvent choisir d'un commun accord ce qu'il paraît pertinent d'intégrer aux cycles à venir. Des options sont par ailleurs proposées, que les pays peuvent refuser, au risque cependant que cela soit relevé par les pays partenaires. La capacité de proposition des référents techniques et des représentants nationaux qui était promise comme étant relativement forte est finalement plutôt faible, du fait du nombre très important de pays participants, et de la diversité des problématiques de chacun. La distinction entre le niveau technique et le niveau stratégique, rend le portage des critiques et revendications techniques difficile et finalement peu arrivent jusqu'au plus haut niveau de décision. Ce fonctionnement est justifié par l'organisation internationale comme permettant à la fois de prendre en compte l'ensemble des avis des pays participants, plusieurs

dizaines, tout en garantissant à la fin une évaluation homogène qui permette de comparer les résultats entre eux. Si en soit le fonctionnement par consensus et par arbitrage est classique des organisations internationales, l'absence de prise en compte des particularismes du système français et le faible pouvoir d'opposition des agents impliqués dans PISA pour la France remet en cause leur rôle d'expert national de la politique éducative française, et des évaluations qui y affèrent.

Cela vient s'insérer dans un contexte politique et administratif français qui ne clarifie par les missions et le rôle alloués à la DEPP par les décideurs politiques. D'un côté, en théorie, la DEPP se voit confirmer au début des années 2000 l'importance de ses missions par l'institutionnalisation de la démarche d'évaluation statistiques avec la mise en place de la LOLF et la mission de représentation et d'expertise auprès de l'OCDE qui lui est allouée. De l'autre, en pratique, elle n'est pas conviée à participer à la construction des indicateurs de pilotage de l'action publique, et son influence auprès de l'OCDE se confirme être très réduite. Plusieurs tentatives sont menées pour retravailler des points de méthode jugés peu adaptés au système français, risquant de rendre les résultats non significatifs. Peu réussissent, y compris en mobilisant d'autres membres des groupes de travail, renforçant encore le sentiment d'une perte de contrôle sur le déroulement de l'évaluation par les évaluateurs nationaux.

Finalement, avant même la publication des premiers résultats, les agents de la DEPP anticipent toute une série de problèmes techniques susceptibles de limiter la validité des résultats de PISA pour la France. L'évaluation est jugée trop proche des méthodes anglo-saxonnes de l'évaluation psychométrique à laquelle les élèves français sont peu habitués. L'investissement important de la DEPP pour des marges d'action faibles en fin de compte est remis en question au sein de la direction, d'autant plus que les équipes ne sont pas renforcées et que les financements ne sont pas complétés à la hauteur des besoins. PISA rogne sur les évaluations nationales, sans que les évaluateurs nationaux ne puissent réellement influencer pour la rendre méthodologiquement adaptée à leurs yeux au système français.

Le rythme très court des publications renforce encore ce sentiment de perte de contrôle sur les résultats de PISA. Alors que les classements ne sont qu'une représentation sans réelle valeur d'analyse, ce sont bien eux qui sont systématiquement repris dans les médias et dans les discours politiques. La DEPP, dont l'une des missions consiste à travailler les résultats PISA de la France pour en tirer des conclusions de politique nationale, n'a pas le temps nécessaire pour la remplir de façon satisfaisante. Le commentaire des résultats n'est pas complété par des conclusions problématisées à visée prospective.

Cette perte d'influence doublée du constat de l'inévitabilité de PISA implique pour la DEPP de redéfinir son champ d'action. Le renforcement de la popularité s'accroît de cycle en cycle, et est renforcé par l'auto-référencement que pratique largement l'OCDE. PISA cite PISA, mais surtout toutes les publications de l'OCDE citent PISA, et pas seulement au sein de la Direction de l'éducation. PISA devient un élément d'évaluation présent en permanence dans les discours, qu'ils traitent de politique éducative, de formation, de politique économique, ou des politiques de l'emploi.

Face à cela, la DEPP s'engage rapidement dans une stratégie visant à consolider son rôle d'expert auprès du politique qui s'appuie sur un discours soulignant la prudence nécessaire dans l'interprétation des résultats de PISA dans le contexte français, qui sont au mieux complémentaires aux expertises nationales. L'évaluation PISA apporte des données complémentaires aux évaluations nationales, qui sont à interpréter avec précaution tant leur construction ne prend pas en compte les particularismes des systèmes éducatifs nationaux. Les classements n'ont en soi pas de sens, et ne sont utiles qu'au bout de plusieurs cycles lorsqu'ils permettent d'évaluer l'évolution des résultats des élèves français par rapport à eux-mêmes dans les cycles précédents. Les discours critiques sur la méthode de PISA laissent rapidement la place à une récupération en interne des concepts centraux de la méthodologie par ailleurs reconnue pour sa qualité technique. La DEPP crée à partir de PISA sa nouvelle évaluation phare, CEDRE, un mélange entre les principes de l'évaluation « à la française » reprenant notamment l'acquisition des programmes scolaires comme objectif de

l'enseignement, mais intégrant des principes méthodologiques clés de PISA comme la théorie de réponse à l'item, et la cyclicité des disciplines étudiées. Elle reprend par ailleurs les résultats PISA à son compte, en publiant davantage de rapports et de notes d'informations fondées sur les résultats de l'évaluation, mettant à jour ce qu'elle considère être des problèmes publics dignes d'intérêt politique, comme le redoublement ou les inégalités sociales et territoriales scolaires. L'objectif est de redonner aux résultats de PISA l'une des fonctions pour lesquelles ils avaient été conçus à l'origine, selon les dires de ses créateurs : l'analyse interne, par chaque pays, de ses résultats aux tests pour éclairer les travaux de leurs instituts de statistique.

Les agents de la DEPP se rapprochent enfin de la recherche, afin de conférer à leurs travaux une plus grande légitimité. Ils ne sont pas seulement des fonctionnaires au service du ministère, mais bien des agents experts à la fois de l'ensemble des méthodes d'évaluation mais aussi des politiques éducatives. Leur mission est d'évaluer mais aussi de rendre compte des résultats dans une démarche prospective, en expliquant par ses données la prégnance de problèmes publics qu'elle souhaite porter à la connaissance du monde universitaire et du monde politique. Elle utilise pour cela notamment sa revue *Education et Formation*, complétée au début des années 2010 par un comité de lecture, et par des sollicitations systématiques d'universitaires reconnus pour leurs travaux sur la question du pilotage des politiques éducatives. Les agents, sous l'impulsion de leurs directeurs et directrices, participent à des colloques, mettant en avant systématiquement leur maîtrise de l'ensemble des méthodes d'évaluation statistiques et leur compréhension de ce qu'elles donnent à voir. Ils deviennent alors les explicateurs de ce pourquoi PISA peut et ne peut pas être utilisé. Cela fait, ils défendent les intérêts de la France comme leur mission l'implique, mais ils défendent aussi leur pratique de l'évaluation et leur expertise d'évaluateur. Au moment où cette expertise se trouve mise à l'épreuve par PISA, tout est fait pour conserver le monopole de la production de savoirs statistiques pour l'orientation de l'action publique, pour garder la maîtrise d'une évaluation qu'on ne maîtrise finalement que peu.

Si cela leur permet de contribuer à la production du savoir, l'engagement de démarches d'évaluation expérimentales ou novatrices reste tout de même limité, mis à part le lancement de CEDRE peut-être. En cela, les agents de la DEPP ne parviennent pas à s'approprier cette expertise créative qui fait le crédit des équipes de l'OCDE et de PISA en particulier. Ces dernières définissent en effet volontiers leur expertise non seulement par leur très haut niveau de spécialisation dans leur domaine, mais aussi par leur maîtrise technique complète des outils de l'évaluation statistique leur permettant créer de nouveaux indicateurs, de nouveaux tests, pour réfléchir à la façon d'évaluer les compétences nécessaires dans un avenir incertain. Les évaluateurs nationaux se trouvent alors contraints d'occuper un rôle de passeur de méthodes et de principes d'évaluation qui sont ceux de PISA, dont ils reconnaissent l'intérêt scientifique mais qui ne correspondent pas à la tradition d'évaluation qu'ils ont construite ces dernières décennies.

Nous avons examiné comment la stratégie de la DEPP a réussi ou non à lui redonner le rôle d'expert évaluateur principal dans la construction de la politique éducative. L'observation des usages différenciés de PISA et des évaluations nationales de la DEPP par les décideurs publics pour dans la fabrique de la loi et le contrôle de l'action de l'État aboutit une conclusion en demi-teinte.

Certes, la mobilisation des évaluations par les parlementaires et les ministres devant la représentation nationale est de plus en plus fréquente. Mais si les travaux de la DEPP, et notamment CEDRE, gagnent en popularité auprès des décideurs politiques, les usages sont monopolisés par PISA. La DEPP vient confirmer les résultats PISA, compenser le retard de publication des données PISA en attendant qu'elles sortent, mais elle ne parvient pas à se voir conférer un rôle bien différent de celui de pourvoyeur de données descriptives d'un état des lieux du système, sans que cela fasse émerger de problème public à traiter. PISA met à l'agenda, tandis que les évaluations de la DEPP, y compris CEDRE, viennent en complément appuyer la bonne qualité des conclusions de l'OCDE. Lorsqu'il s'agit d'exiger une réforme, d'alerter sur un point critique, PISA est

l'évaluation venant appuyer les argumentaires que ce soit des parlementaires ou des membres du gouvernement.

En majorité les usages de PISA sont par ailleurs relatifs au classement de la France parmi les autres pays participants. La DEPP n'est pas réellement parvenue à changer le prisme d'analyse de PISA. Même après plusieurs cycles, les classements sont repris très régulièrement par les acteurs politiques au moment d'utiliser PISA. Cela témoigne de l'importance croissante du regard international sur la façon dont sont menées les politiques éducatives nationales. Portée par les différentes stratégies européennes d'intégration des systèmes éducatifs et de formation depuis le début du XXe siècle, l'évaluation ne rend plus seulement compte de l'action de l'Etat devant les citoyens ou devant la représentation nationale. PISA plus que toute autre implique que l'on rende compte de son action nationale aux partenaires internationaux, et plus seulement européens. Elle implique également que l'on rende compte à l'OCDE, qui développe par ailleurs dans les dernières années de nouveaux « produits » intégrant les résultats PISA, qui concluent sur les « bonnes réformes » et ce que les décideurs politiques font qui va dans le « bon sens ». L'usage abondant de PISA a bien eu pour effet de faire émerger un nouvel œil critique des actions menées par les décideurs politiques français en matière d'éducation : l'organisation qui pilote l'évaluation, l'OCDE. Si les résultats PISA fondent une réforme, ils ont légitimement le rôle d'indicateur du succès de cette réforme. La DEPP perd alors le contrôle d'une partie des indicateurs de résultat de l'action éducative nationale, dont elle a la responsabilité. Avec cela s'étiole la définition de la performance scolaire comme la capacité des élèves à maîtriser des connaissances académiques constituées en programmes scolaires, pour s'orienter davantage vers l'acquisition de compétences à mobiliser au quotidien.

A la fin de notre période, PISA constitue un point de référence clé dans la construction de la loi et dans le contrôle de l'action du gouvernement par les parlementaires. Le benchmark proposé par l'OCDE permet en quelque sorte de dépasser la limite de la réalisation d'évaluations *ex post*, une fois les politiques mises en place. Les résultats PISA donnent à voir les résultats potentiels de

réorientations de politique éducative, bien que les contextes de réalisation soient différents. Ils sont employés à plusieurs reprises pour renforcer la légitimité des réformes auprès des citoyens et des élus. Si en 2005 François Fillon avait déjà utilisé PISA pour lier la loi pour l'avenir de l'école au besoin d'inscrire l'acquisition d'un socle commun de compétence comme objectif de l'école, la légitimation d'une réforme par PISA est encore plus évidente en 2013 avec la loi Peillon qui est justifiée par les recommandations de l'OCDE de lutter contre les inégalités scolaires en France. Si l'attrait des élus à droite de l'échiquier politique pour les conclusions de l'OCDE était alors bien connu, son appropriation par les familles de gauche n'était pas acquise. La réforme de 2013 marque un vrai tournant de ce point de vue. PISA ne change certes pas réellement les propositions des acteurs politiques toujours clivés entre droite et gauche, quitte à réinterpréter les indicateurs de l'OCDE. En revanche, son adoption comme registre argumentatif par tous, y compris les groupes politiques traditionnellement critiques de l'OCDE archétype de l'idéologie néolibérale, marque un changement de façon de concevoir l'évaluation de l'action de l'État, l'identification de problèmes publics et la préparation des réformes jugées souhaitables. PISA s'est en quelque sorte banalisé et devient un outil légitime d'évaluation de l'école sans distinction de projet politique. Chacun reprenant à son compte les indicateurs qui permettent de soutenir sa propre ligne politique. C'est là véritablement la force de la publication de l'OCDE. La grande diversité d'indicateurs qu'elle propose permet à chacun d'y trouver les mots-clés qui lui conviennent. De manière fort peu surprenante, les questions se rapportant aux inégalités scolaires sont reprises rapidement par les parlementaires de gauche. Les parlementaires de droite traiteront davantage des questions de financement du système éducatif, trop coûteux en comparaison de ceux des partenaires européens et occidentaux et aux résultats moins bons lorsque l'on regarde PISA. Le résultat est là : PISA est devenu la principale source d'information sur le fonctionnement du système scolaire et ses effets pour les élus et responsables politiques nationaux.

L'accentuation des usages de l'évaluation, visible dans toutes les familles politiques, a pour dommage collatéral de raviver les débats de la fin du XXe siècle

sur l'indépendance d'une expertise d'évaluation interne au ministère. La question est posée : quel degré d'indépendance est acceptable pour que l'évaluation soit considérée comme légitime et ses résultats valables ? A chaque réforme se pose la question de l'indépendance de la DEPP compte tenu de sa proximité avec le politique en étant intégrée au ministère. Le CNESCO est même créé à l'occasion de la réforme de 2013, pour évaluer les résultats du système éducatif. Les données de la DEPP risqueraient d'être orientées pour convenir au ministre, voire cachées pour celles qui porteraient préjudice à l'action du gouvernement en place. Cela influe très certainement sur la façon dont les données de la DEPP sont utilisées dans le cadre des travaux parlementaires, davantage pour décrire que pour expliquer. L'appréciation de l'indépendance de la DEPP est cependant tout à fait subjective : on craint l'accointance ou la pression politique des évaluations lorsqu'elles sont menées en propre, mais on ne remet pas en cause les résultats d'une évaluation internationale dans laquelle les agents du ministère sont supposés influencer pour que soient pris en compte les particularités du système français – et qu'ils mettent eux-mêmes concrètement en œuvre. Alors que l'OCDE valorise la légitimité de son évaluation par sa construction au consensus des pays participants, les décideurs politiques de droite comme de gauche, de la majorité ou de l'opposition, légitiment PISA au contraire par l'effacement de la participation notamment de la DEPP.

Il ne faut pas cependant conclure à la perte irrémédiable de légitimité de la DEPP comme évaluatrice de l'action éducative. Elle est bel et bien sollicitée pour sa maîtrise des évaluations nationales, comme des évaluations internationales, et joue le rôle de transmetteur de savoir sur toutes les évaluations sur lesquelles elle travaille, y compris PISA. Elle continue par ailleurs d'entreprendre son travail de valorisation de l'indépendance de son programme de travail qui porte ses fruits dans le monde universitaire, en témoigne le nombre croissant de colloques auxquels les agents de la DEPP, notamment ceux chargés de PISA, sont invités à intervenir.

Par ailleurs, et cela mériterait sans doute que des travaux s'y intéressent tout particulièrement, la DEPP continue d'étendre son public cible au-delà des

décideurs politiques et des acteurs de la recherche, en se positionnant comme passeur des résultats PISA auprès des enseignants. Ces derniers ne sont a priori pas les cibles attendues des conclusions de l'OCDE. L'évaluation n'évalue pas en tant que telle l'action d'enseignement, mais bien les résultats du système éducatif. Les recommandations issues des résultats ne traitent pas des pratiques pédagogiques à favoriser dans les classes, tout au plus ont elles incité à revoir la place accordée à la formation pédagogique des enseignants en France. Cela transparaît clairement dans les discours des interlocuteurs de la DEPP, de l'OCDE et du ministère que nous avons pu rencontrer. Il n'est pas question de faire du « PISA like », c'est-à-dire de prévoir dans les consignes données aux enseignants de travailler pour assurer de meilleurs résultats à PISA. Néanmoins, force est de constater que les méthodes et les items de PISA ont parcouru un chemin menant jusqu'aux salles de classe. Certains concepts clés des programmes, comme en mathématiques les « tâches complexes », ont été inspirés de ce que pratique PISA. Les examens de fin de cycle, comme le brevet ou le baccalauréat, ont également été sujets à des réformes dans les années 2010 et ont mis en pratique des modes d'évaluation proches pour certains de ce que fait PISA. Charge aux enseignants de préparer ensuite leurs élèves à passer ces examens, avec des méthodes permettant donc, en corollaire et pour ceux qui seraient amenés à participer à PISA, d'améliorer leurs résultats. Enfin, les manuels de formation des enseignants et les modules dispensés dans les ESPE à partir de leur création ont également utilisés certains items libérés de PISA. Dans ce mouvement, la DEPP s'est rapprochée en fin de période des enseignants, des formateurs et des inspections générales afin de porter les éléments d'analyse qu'elle estime pertinente de partager au monde enseignant. Si cela a été largement lié au départ à la personnalité de Catherine Moisan, sa successeuse Fabienne Rosenwald et les responsables de la DEPP impliqués dans PISA ont perpétué cette diffusion d'information, faisant des enseignants de nouvelles cibles des résultats PISA. Il n'a pas été possible de davantage traiter ce nouveau rôle auto-défini de la DEPP, qui gagnerait à être étudié de façon plus approfondie pour

comprendre ce que cela implique à la fois pour l'expertise de la DEPP et pour l'appropriation de PISA par le corps enseignant.

Pour finir, on peut souligner que deux autres sujets paraissent mériter par ailleurs un développement qui n'a pas pu être mené ici. Le temps du doctorat, bien que long, a nécessité de clôturer les recherches de terrain et documentaires au moment de l'émergence de nouvelles problématiques issues aussi bien des changements ministériels survenus après l'élection d'Emmanuel Macron en 2017, après le changement de direction de la DEPP et le remplacement de Catherine Moisan par Fabienne Rosenwald fin 2015, et l'événement international politique de l'année 2020 qu'a été le démarrage de la crise du covid-19.

D'abord, on l'a vu, les appropriations par les ministres des évaluations sont représentatives du rôle alloué à la DEPP dans l'accompagnement du pilotage de la politique éducative. La nouvelle posture de Jean Michel Blanquer sur l'importance des évaluations nationales des résultats des élèves comme outil déterminant du déclenchement de dispositifs qu'il a développé à plusieurs moments de la scolarité, notamment l'accompagnement contre le décrochage en école primaire et en début de collège, marque sans doute le début d'une nouvelle étape dans les missions de la DEPP. La direction se voit confier le soin de mener, en plus du reste de ses activités, les évaluations de début de CP et de CE1 et d'évaluer les résultats des élèves à la suite des mesures prises par le ministre comme le dédoublement des classes de primaire en Zone d'Education Prioritaire. L'évaluation PISA apparaît dans ce cadre moins présente dans le discours du ministre. Celui-ci redonne en revanche aux évaluations nationales « diagnostics » un rôle à la fois d'éclairage des résultats des mesures prises par le ministère (analyse des résultats anonymisés des élèves) et d'accompagnement à la scolarité individualisé puisque chaque élève devra recevoir un retour sur ses résultats. L'arrivée en 2017 de Jean-Michel Blanquer à la tête du ministère de l'Éducation nationale aura été trop tardive pour être intégrée dans nos travaux. Il sera intéressant de poursuivre l'observation du repositionnement de l'expertise de la

DEPP à l'aune de ce changement de posture ministérielle sur les objets et les objectifs de l'évaluation nationale des résultats des élèves.

Enfin, la dernière évolution récente porte sur la position de l'OCDE dans le champ des experts en politiques éducatives. L'organisation internationale a amorcé un clair développement de son activité de conseil aux États et de réaction aux situations de crise, en particulier dans la situation internationale exceptionnelle de l'année 2020. Tout au long de la crise sanitaire liée au covid 19 d'abord, l'OCDE a travaillé sur la publication de notes et de recommandations portant sur ce que PISA dit de la préparation des systèmes éducatifs à ce type de crise et à leur capacité à garantir l'accès à l'éducation de leurs élèves. L'organisation s'inscrit ici dans une conception de l'expertise en politique publique en contexte de crise, impliquant une réaction d'urgence en termes de politiques de santé ou de politiques environnementales. Les évaluations PISA, par le nombre important de thématiques qu'elles recouvrent et plus seulement les résultats secs des élèves aux tests, a permis à l'organisation de compenser le report de l'évaluation PISA 2023 par la publication de ces notes rendant compte des fractures numériques, de la capacité de développer l'enseignement à distance, du ressenti des chefs d'établissements sur leur capacité à adapter les enseignements à la situation de crise et notamment aux confinements, des risques de renforcement des inégalités scolaires du fait des pertes d'apprentissage liées à la crise. Un questionnaire porté auprès des chefs d'établissement dans le cadre de PISA 2018 avait notamment largement abordé la capacité d'inventer et de mener des pratiques d'apprentissage à distance dans leur établissement. Cela a mis en évidence deux choses. D'un côté, sur le fond, PISA semble avoir rempli sa mission d'anticipation des sujets d'intérêt à l'avenir dans le pilotage de l'action éducative. Bien évidemment elle n'a pas traité le sujet sous l'angle de la gestion de crise sanitaire ni de crise internationale. En revanche, elle a permis d'adapter le contenu et le format des publications pour devenir expert de la gestion de crise en fournissant des informations sur les capacités de chacun à développer dans l'immédiat un enseignement numérisé et les risques qui y sont relatifs notamment en termes d'égalité d'accès à une continuité éducative. Sur un second plan, PISA

a montré la capacité à rester un objet constant de discussion et d'intérêt pour le pilotage de l'action éducative, même lorsqu'elle n'était pas en capacité de produire ce qui est effectivement attendu d'elle de façon prioritaire – une évaluation cyclique tous les trois ans. L'impossibilité de mener les phases de test pour PISA 2021 a amené l'organisation à repousser le prochain cycle PISA à 2023. Elle a réussi néanmoins à maintenir sa régularité de publication et à demeurer un objet d'intérêt central pour les décideurs politiques des pays participants sans créer de nouvelles enquêtes *ad hoc* sur la période de crise sanitaire. Ce n'est pas le cas de la DEPP qui a lancé sept enquêtes en mai 2020, auprès à la fois d'élèves du second degré et de leurs parents, de professeurs d'écoles, de collèges et lycées et de personnels de l'Éducation nationale. En tout, près de 100 000 personnes ont été interrogées spécifiquement sur l'expérience scolaire lors des confinements, l'autonomie des élèves, le contact entre les élèves et les professeurs. Les prochaines années nous diront si ces nouvelles manifestations d'expertise auront pour effet de repositionner encore l'expertise des évaluateurs nationaux et internationaux vers un pilotage continu de l'action éducative complétée d'une capacité à émettre des données utiles pour les décideurs politiques en gestion de crise.

Bibliographie

Ouvrages

Ouvrages complets

Abbott Andrew, *The System of Professions, an Essay on the Division of Expert Labor*, University of Chicago Press, 1988.

Anderson James E. , *Public Policymaking*, Praeger, 1975.

Austin John L., *How to Do Things with Words*, edited by J.O. Urmson and M. Sbisà, Oxford : Oxford University Press, 2nd edition : 1976 ; traduction française de G. Lane, *Quand dire c'est faire*, Paris : Editions du Seuil, 1970 ; réédité avec une postface de F.Récanati dans la collection « Points-essais », 1991.

Baker Frank B. et Kim Seock-Ho, *Item Response Theory – Parameter Estimation Techniques*, 2e ed. New York, Marcel Dekker, 2004.

Baudelot Christian, Establet Roger, *L'école capitaliste en France*, Paris, Maspero, 1971.

Baudelot Christian, Establet Roger, *L'école primaire divisée*, Paris, Maspero, 1979.

Baudelot Christian, Establet Roger, *Le niveau monte*, Paris, Le Seuil, 1988.

Baudelot Christian, Establet Roger, *Allez les filles*, Paris, Le Seuil, 1992.

Baudelot Christian, Establet Roger, *Quoi de neuf chez les filles ? Entre stéréotypes et libertés*, Nathan, coll. L'enfance en questions, 2007.

Baudelot Christian et Establet Roger, *L'élitisme républicain : l'école française à l'épreuve des comparaisons internationales*. Paris, Ed du seuil, 2009.

Becker Gary, *Human Capital : A Theoretical and Empirical Analysis, with Special Reference to Education*, University of Chicago Press, 1994.

Becker Howard, *Outsiders : Studies in sociology of deviance*, Free Press, 1966.

Bertrand Richard, Blais Jean-Guy, *Modèles de mesure. L'apport de la théorie des réponses aux items*, Sainte-Foy, Presses de l'Université du Québec, 2004.

Bezes Philippe, *Réinventer l'Etat, les réformes de l'administration française (1962-2008)*, Paris, PUF, 2009.

Blum Alain et Guérin-Pace France, *Des lettres et des chiffres : des tests d'intelligence à l'évaluation du « savoir lire », un siècle de polémiques*, Paris, Fayard, 2000.

Bonnet Gérard et al., *The Use of National Reading Test for International Comparisons : Ways of Overcoming Cultural Bias*. Paris : Ministère de l'Éducation nationale, Direction de la Programmation et du Développement, 2001.

Bourdieu Pierre, Passeron Jean-Claude, *Les Héritiers, les étudiants et la culture*, Paris, Editions de Minuit, 1964.

Bourdieu Pierre, Passeron Jean-Claude, *La Reproduction*, Paris, Editions de Minuit, 1970.

Boussaguet, Laurie, Jacquot Sophie et Ravinet Pauline, *Dictionnaire des politiques publiques*. 4^e édition précédée d'un nouvel avant-propos. Presses de Sciences Po, 2014.

Bruno Isabelle et Didier Emmanuel, *Benchmarking. L'État sous pression statistique*, Paris, La Découverte, 2013.

Citron Suzanne, *L'école bloquée*, Paris, Bordas, 1971.

Crozier Michel et Friedberg Ehrard, *L'acteur et le système*, Paris, Ed. du Seuil, 1977.

Data Lorraine, *Le grand trucage. Comment le gouvernement manipule les statistiques*, Paris, La Découverte, 2009.

Delasalle Marine, Georgakakis Didier (dir.), *The Political Use of Governance – Studying an EU White Paper*, Toronto, Barbara Budrich Publishers, 2012.

Delmas Corinne, *Sociologie politique de l'expertise*, Paris, La Découverte, 2011.

Derrida Jacques, *Limited Inc.*, Galilée, Paris, 1990.

Desrosières Alain, *Gouverner par les nombres*, Presses des mines, 2008.

Dominique Julia, *Les trois couleurs du tableau noir*, Paris, Belin, 1981.

Dubar Claude, Tripier Pierre et Boussard Valérie, *Sociologie des professions*, Paris, Armand Collin, 2011.

Dubran Patrice, *Penser l'action publique*, Paris, LGDJ, 2010.

Dumoulin Laurence, La Branche Stéphane, Robert Cécile et Warin Philippe, *Le recours aux experts : raisons et usages politiques*, Presses universitaires de Grenoble, Grenoble, 2005.

Geay Bertrand, *Le syndicalisme enseignant*, Paris, La Découverte, 2005.

Gerstlé Jacques, *La communication politique*, Armand Colin, Paris, ed. 2011.

Gurgand Marc, *Economie de l'éducation*, La Découverte, Paris, 2005.

Henry Emmanuel, *Amiante : un scandale improbable, sociologie d'un problème public*, Presses universitaires de Rennes, Rennes, 2007.

Henry Emmanuel, « Nouvelles dynamiques de savoirs et permanence des rapports de pouvoir. L'impact – limité – des transformations – importantes – de l'expertise en santé au travail », *Revue française de science politique*, vol. 61, n°4, 2011.

Henry Emmanuel, et al., *Dictionnaire critique de l'expertise. Santé, travail, environnement*. Presses de Sciences Po, 2015.

Iriye Akira, *Global Community: the Role of International Organizations in the Making of the Contemporary World*, University of California Press, Berkeley, 2002.

Isambert-Jamati Viviane, *Les savoirs scolaires : enjeux sociaux et contenus d'enseignement et de leurs réformes*, Paris, Ed. Universitaires, 1990.

Kingdon John, *Agendas, Alternatives and Public Policies*, New-York, Harper Collins, 1984.

Klinkenberg Jean-Marie, *La langue et le citoyen. Pour une politique de la langue française*, Paris, Presses universitaires de France, 2001.

Lagroye Jacques, François Bastien, Sawicki Frédéric, *Sociologie politique*, Paris, Presses de Sciences-Po – Dalloz, 2002.

Lascombes Pierre et Le Galès Patrick, *Gouverner par les instruments*, Paris, Presses de Sciences Po, 2005.

Lascombes Pierre et Le Galès Patrick, *Sociologie de l'action publique*, Paris, Armand Colin, 2009.

Laveault Dany et Grégoire Jacques, *Introduction aux théories des tests en psychologie et en sciences de l'éducation*, Bruxelles, De Boeck, 2002.

Leclercq Dieudonné, *La conception des questions à choix multiples*, Bruxelles, Labor, 1986.

Leclercq Dieudonné, *Qualité des questions et signification des scores*, Bruxelles, Labor, 1987.

Mangez Eric, *Réformer les contenus d'enseignement*, PUF, Paris, 2008.

Martens Kerstin, Rusconi Alessandra et Leuze Kathrin, *New Arenas of Education Governance. The Impact of International Organizations and Markets on Education Policy Making*, Palgrave, 2007.

Maugeri Salvatore (dir.), *Délit de gestion*, La Dispute, Paris, 2001.

Merrien François-Xavier, *L'Etat-providence*, Paris, PUF coll « Que sais-je », 2007 [3e éd.].

Neveu Erik, *Sociologie politique des problèmes publics*, Armand Colin, 2015.

Normand Romuald, *Gouverner la réussite scolaire. Une arithmétique politique des inégalités*, Berne : Peter Lang et Lyon : ENS de Lyon, 2011.

Nye Joseph, *Bound to lead, The changing nature of American power*, Basic Book, New York, 1990.

Pélisse Jérôme, Protais Caroline, Larchet Keltoume, Charrier Emmanuel (dir.), *Des chiffres, des maux et des lettres. Les experts judiciaires en économie, en psychiatrie et en traduction*, Paris, Armand Colin, coll. « Recherches », 2012.

Plantey Alain, *Principes et développement du processus international de décision, L'avenir des organisations internationales*, Paris, 1983.

Pons Xavier, *Evaluer l'action éducative, des professionnels en concurrence*, Education et société, PUF, Paris, 2010.

Prost Antoine, *Histoire de l'enseignement en France, 1800-1967*, Paris, Armand Colin, 1968.

Prost Antoine, *Education, société et politiques. Une histoire de l'enseignement en France de 1945 à nos jours*, Paris, Le Seuil, 1992.

Prost Antoine, *Du changement dans l'école. Les réformes de l'éducation de 1936 à nos jours*, Paris, Seuil, 2013.

Thélot Claude, *Tel père, tel fils ?*, Paris, Dunod, Bordas, 1982.

Thélot Claude, *L'évaluation du système éducatif*, Paris, Nathan, 1993.

Thélot Claude, *Les écrivains racontent l'école*, Delagrave, 2001.

Thélot Claude, *L'origine des génies*, Paris, Le Seuil, 2003.

Thélot Claude et Marchand Olivier, *Le travail en France, 1800-2000*, Nathan, collection Essais et Recherches, Série Sciences Sociales, Paris, 1997.

Thélot Claude et Joutard Philippe, *Réussir l'école*, Paris, Editions du Seuil, 1999.

Toulemonde Bernard, *Petite histoire d'un grand ministère*, Paris, Albin Michel, 1988.

Trépos Jean-Yves, *La sociologie de l'expertise*, PUF, Que-sais-je ?, 1996.

Vial Monique, *Les débuts de l'enseignement spécial en France*, Travaux du CRESAS, 1979, n°18.

Weber Max, *Le savant et le politique*, nouvelle traduction de Catherine Colliot-Thélène, Paris, La découverte, 2003.

Chapitres d'ouvrages et notices

Bérard Yann et Crespin Renaud, « Situation d'expertise », Emmanuel Henry éd., *Dictionnaire critique de l'expertise. Santé, travail, environnement*. Presses de Sciences Po, 2015.

Bezes Philippe, « Pour une sociologie des politiques de réforme de l'administration », dans *Réinventer l'Etat, les réformes de l'administration française (1962-2008)*, Paris, PUF, 2009.

Braud Philippe, « L'étude des acteurs », Philippe Braud éd., *La science politique*. Presses Universitaires de France, Paris, 2017.

Felouzis Georges, Charmillot Samuel, « Comment sont construites les enquêtes PISA ? », dans : Georges Felouzis éd., *Les enquêtes PISA*, Presses Universitaires de France, « Que sais-je ? », 2012

Forestier Christian, « Le meilleur système du monde ? » dans Jean-Claude Emin, Christian Forestier, Claude Thélot, *Que vaut l'enseignement en France ?*, Editions Stock, 2001.

Geay Bertrand, « Le « tous ensemble » des enseignants », dans Sophie Bérout et René Mouriaux, *L'année sociale*, Paris, Syllepse, 2004.

Grek Sotiria, « Educational Governance and Transnational Experts. Europe and Its Emerging Relationship with the OECD », dans Martin Lawn, Romuald Normand, *Shaping of European Education. Interdisciplinary Approaches*, Abingdon, Routledge, 2015.

Hambleton Ronald K. et Li Shuhong, « Translation and Adaptation Issues and Methods for Educational and Psychological Tests », dans Frisby Craig L. et Reynolds Cecil R. (Eds.), *Comprehensive handbook of multicultural school psychology*, John Wiley & Sons, 2005.

Leclerc Olivier, « L'indépendance de l'expert ». K. Favro (coord.). *L'expertise : enjeux et pratiques*, Lavoisier/Tec&Doc, Sciences du risque et du danger, 2009,

Maillard Jacques et Kübler Daniel, « Chapitre 1. Les processus de mise à l'agenda », *Analyser les politiques publiques*, sous la direction de Jacques Maillard et Daniel Kübler, Presses universitaires de Grenoble, 2016.

Maroy Christian, « Politiques et outils de « l'école de la performance » : accountability, régulation par les résultats et pilotage. » Dans C. Maroy (dir.), *L'école à l'épreuve de la performance : les politiques de régulation par les résultats* (pp. 14-31). Bruxelles : De Boeck, 2013.

Mottier Lopez Lucie, « L'évaluation en éducation : des tensions aux controverses. » Dans Lucie Mottier Lopez et M. Crahay (dir.), *Évaluations en tension : entre la régulation des apprentissages et le pilotage des systèmes* (pp. 7-25). Bruxelles : De Boeck, 2009.

Palier Bruno, « Chapitre 7 : Les instruments, traceurs du changement. La politique des retraites en France », Pierre Lascoumes éd., *Gouverner par les instruments*. Paris, Presses de Sciences Po, 2005.

Peters B. Guy, « Nouveau management public (*new public management*) », Laurie Boussaguet éd., *Dictionnaire des politiques publiques. 4^e édition précédée d'un nouvel avant-propos*. Presses de Sciences Po, 2014.

Plantey Alain, Lorient François, « Chapitre 1. Nature et champ de la fonction publique internationale », *Fonction publique internationale: Organisations mondiale et européennes*. Paris; CNRS Editions, 2005.

Polidano Charles, « Administrative reform in core civil services : application and applicability of the new public management » chapitre dans Willy Mc Court et Martin Minogue (ed.), *The Internationalization of Public Management*, chapter 3, Edward Elgar Publishing. 2001

Rizvi Fazal, Lingard Bob, « The OECD and Global Shifts, in Education Policy », dans Robert Cowen & Andreas M. Kazamias (ed.), *International Handbook of Comparative Education*, Dordrecht : Springer Netherlands, 2009.

Robert Cécile, « Expertise et action publique », dans Olivier Borraz, Virginie Guiraudon (dir.), *Politiques publiques. 1, La France dans la gouvernance européenne*, Paris, Presses de Sciences Po, 2008.

Thélot Claude, préface de l'ouvrage de Jean-Claude Emin, Christian Forestier, Claude Thélot, *Que vaut l'enseignement en France ?*, Editions Stock, 2001.

Thèses

Bellarbre Elodie, *Compenser les inégalités dans le secondaire en France : études à partir des enquêtes PISA, CEDRE et TALIS*, thèse pour le doctorat d'économie sous la direction de Jean-François Giret, soutenue à l'Université de Bourgogne Franche Comté le 19 octobre 2017.

Catteau Damien, *La LOLF et la modernisation de la gestion publique : la performance, fondement d'un droit public financier rénové*, thèse pour le doctorat de droit public financier sous la direction de Xavier Vandendriessche, soutenue à l'Université de Lille 2 en 2005.

Clément Pierre, *Réformer les programmes pour changer l'école. Une sociologie historique du champ du pouvoir scolaire*, thèse pour le doctorat de sociologie sous la direction de Bertrand Geay et soutenue à l'Université de Picardie Jules Verne le 11 mars 2013.

Gaillard Brice, *La performance de l'action publique territoriale*, thèse pour le doctorat de science politique sous la direction de Pascal Jan, soutenue à l'Université de Bordeaux le 16 décembre 2016.

Gayon Vincent, *L'OCDE au travail. Contribution à une sociologie historique de la « coopération économique internationale » sur le chômage et l'emploi (1970-2010)*, thèse pour le doctorat en science politique sous la direction de Brigitte Gaïti, soutenue le 30 novembre 2010 à l'Université Paris Dauphine.

Simha Jules, *Les expérimentations sociales en France. Une sociologie de l'évaluation des politiques publiques*, thèse pour le doctorat de sociologie sous la direction d'Annette Jobert, soutenue le 3 décembre 2015 à l'Université Paris Ouest Nanterre La Défense.

Articles

Articles de revue

Adler E., Haas P. M., Conclusion: Epistemic Communities, World Order, and the Creation of a Reflective Research Program, *International Organization*, 46/1, 1992.

Afonso Natércio et Costa Estela, *Use and Circulation of OECS's Programme for International Student Assessment (PISA) in Portugal*, Rapport pour le projet Know&Pol, Université de Lisbonne, Faculté de psychologie et de sciences de l'éducation, 2009.

Aggeri Franck, « Qu'est ce que la performativité peut apporter aux recherches en management et sur les organisations. Mise en perspective théorique et cadre d'analyse », *M@n@gement*, 1, vol.20, 2017.

Akrich Madeleine, « From Communities of Practice to Epistemic Communities: Health Mobilizations on the Internet », *Sociological Research Online*, 15/2, 2010.

Ambroise Bruno, *Les conditions de la performativité et l'efficacité des actes de parole – réponse à J. Butler*, Poitiers, Mars 2007.

Ammermüller Andreas, « PISA : What Makes the Difference ? », *Explaining the gap in PISA test scores between Finland and Germany*, ZEW Discussion Papers, n° 04-04, Centre for European Economic Research, Mannheim, 2004.

Arveiller Jacques, « La loi de 1909 et la définition de l'instituteur spécialisé », *Les Sciences de l'éducation - Pour l'Ère nouvelle*, 1, Vol. 42, 2009.

Avvisati Francesco et Keslair François, « REPEST, Stata module to run estimations with weighted replicate samples and plausible values », *Statistical Software Components*, Boston College Department of Economics, 2020.

Backouche Isabelle, « Introduction », numéro spécial Devenir expert, *Genèses*, 70, 1, 2008.

Barbier Jean-Claude, « Eléments pour une sociologie de l'évaluation des politiques publiques en France », *Revue française des affaires sociales*, n°1-2, 2010.

Baye Ariane, « La gestion des spécificités linguistiques et culturelles dans les évaluations internationales de la lecture », *Politiques d'éducation et de formation, analyses et comparaisons internationales*, Cecoeduc, vol. 11, 2004.

Béland Sébastien, Magis David et Raïche Gilles, « Estimation des paramètres d'item et de sujet à partir du modèle de Rasch : une étude comparative des logiciels BILOG-MG, ICL et R ». *Mesure et évaluation en éducation*, 36(1), 2013.

Belorgey Nicolas, « Les cadres de l'expertise publique. Les trois dimensions de l'expertise d'État saisies par une ethnographie des économistes de la Banque de France », *Genèses*, 85, 2011.

Benamouzig Daniel, Besançon Julien, « Administrer un monde incertain : les nouvelles bureaucraties techniques. Le cas des agences sanitaires en France », *Sociologie du travail*, 47 (3), 2005.

Benamouzig Daniel, Barbier Jean-Claude, Geoffard Pierre-Yves et Duvoux Nicolas, « Les sciences sociales à l'épreuve de l'expertise », *Sociologie*, n°1, vol.2, 2011.

Berrebi-Hoffmann Isabelle et Grémion Pierre, « A quoi servent les experts ? », *Cahiers internationaux de sociologie*, n°126, 2009.

Bessonneau Pascal, Arzoumanian Philippe et Pastor Jean-Marc, « Une évaluation sous format numérique est-elle comparable à une évaluation de type « papier-crayon » ? », *Education & Formations*, n°86-87, mai 2015.

Bezes Philippe, « Aux origines des politiques de réforme administrative sous la Ve République : la construction du « souci de soi de l'Etat » », *Revue Française d'administration publique*, 2, n°102, 2002.

Bezes Philippe, « Le renouveau du contrôle des bureaucraties. L'impact du New Public Management », *Informations sociales*, vol. 126, n°6, 2005.

Bezes Philippe, Demazière Didier, Le Bianic Thomas, Paradeise Catherine, Normand Romuald, Benamouzig Daniel, Pierru Frédéric et Evetts Julia, « New Public Management et professions dans l'État : au-delà des oppositions, quelles recompositions ? », *Sociologie du travail*, vol.53, n°3, 2011.

Brunetière Jean-René, « Les indicateurs de la loi organique relative aux lois de finance (Lolf) : une occasion de débat démocratique ? » *Revue française d'administration publique*, n°117, 2006.

Binet Alfred et Simon Théodore, « Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux », dans *L'année psychologique*, vol. 11, 1904.

Boisivon Jean Pierre, « L'horizon de l'école », *Projet*, n°223, automne 1990.

Bonnet Gérard, « Vers une méthode alternative pour l'évaluation de la lecture dans les enquêtes internationales », *Education & formations* n°70, décembre 2004.

Bosk Charles et Hilgartner Stephen, « The Rise and Fall of Social Problems : A Public Arena Model », *American Journal of Sociology*, vol.94, n°1, 1988.

Bottini Fabien, « Le new public management : facteur d'un alignement des fonctions publiques américaine et française sur le droit commun du travail ? », *Les politiques du marché du travail face à la crise économique*, 2014.

Bottini Fabien, « L'impact du new public management sur la réforme territoriale », *Revue française de droit administratif*, Dalloz, 2015.

Bouvet Renaud, « L'indépendance de l'expert », *Revue Juridique de l'Ouest*, Année 2015.

Brunetière Jean-René, « Les indicateurs de la loi organique relative aux lois de finances (LOLF): une occasion de débat démocratique? », *Revue française d'administration publique*, Ecole nationale d'administration, n°117, 2006.

Bruno Isabelle, « La déroute du « benchmarking social ». La coordination des luttes nationales contre l'exclusion et la pauvreté en Europe », *Revue Française de Socio-Economie*, La Découverte, n°5, 2010.

Busch Per Olof et Jorgens Helge, « The International sources of policy convergence : explaining the spread of environmental policy innovations », *Journal of european public policy*, 12 octobre 2005.

Caille Jean-Paul, « Quarante-cinq ans de panels d'élèves à la Direction de l'évaluation, de la prospective et de la performance (DEPP) », *Education et Formations*, n°95, DEPP, MEN, 2017.

Calafat Guillaume, « Expertise et compétences: procédures, contextes et situations de légitimation », *Hypothèses*, 1, 14, 2011.

Catteau Damien, *La LOLF et la modernisation de la gestion publique : la performance, fondement d'un droit public financier rénové*, thèse de doctorat sous la direction de Xavier Vandendriessche, 2005.

Chatel Elisabeth, « L'évaluation de l'éducation dans le contexte de la mondialisation : le cas français », *Tréma*, 40, 2013.

Chevallier Jacques, « L'entrée en expertise », *Politix*, vol.9, n°36, quatrième trimestre, 1996.

Coppé Sylvie et Houdement Catherine, « Réflexion sur les activités concernant la résolution de problème à l'école primaire ». *Grand N*, 69, 2002.

Coulet Jean-Claude, « Les notions de compétences clés : l'éclairage d'un modèle théorique fondé sur l'analyse de l'activité », *Activités*, 13-1, 2016.

Cognet Georges, « Les trois vies du Binet-Simon : Binet-Simon, Nemi, Nemi-2 », *Recherches et éducations*, octobre 2011.

Crahay Marcel et Delhaxhe Arlette, « L'analyse comparée des systèmes éducatifs : entre universalisme et particularisme culturel », Jean-Paul Bronckart éd., *Transformer l'école*, De Boeck Supérieur, 2004.

Cytermann Jean-Richard, « Evaluer et prévoir. De la DEP à la DPD, *Perspectives documentaires en éducation*, n°54, 2001.

Cytermann Jean-Richard, « La contribution des outils statistiques et d'évaluation à la modernisation de l'Education nationale », *Politiques et management public*, vol.23, n°1, 2005.

Cytermann Jean-Richard, « Le financement du système éducatif français et l'enjeu de la performance », *Revue internationale d'éducation de Sèvres*, n°65, avril 2014.

Darby Michael et Karni Edi, « Free competition and the optimal amount of fraud », *Journal of Law and Economics*, vol. 16, 1973.

Dakowska Dorota, « What(ever) works : les organisations internationales et les usages de « bonnes pratiques » dans le domaine de l'enseignement supérieur », *Critique Internationale*, Presses de sciences po, 2017.

Dakowska Dorota, « Créer des experts à son image : la Commission européenne et les politiques de l'enseignement supérieur », *Politix*, De Boeck Supérieur, 2020.

De Bouttemont Cécile, « Etats-Unis : les tests d'évaluation en débat », *Revue internationale d'éducation de Sèvres*, n°53, avril 2010.

De Padirac Hortense et Rozenberg Olivier, « L'évaluation au Parlement français : l'heure des choix », *Sciences Po LIEPP Policy Brief*, 2019.

De Padirac Hortense, « Le Parlement français et l'évaluation : une institutionnalisation impossible ? », *Sciences Po LIEPP*, working paper n°80, octobre 2018.

Defresne Florence et Krop Jérôme, « La massification scolaire sous la Ve république, une mise en perspective des statistiques de l'Education nationale (1958-2014) », *Education et formations*, n°91, septembre 2016.

Demeestere René et Orange Gérald, « Gestion publique : qu'est-ce qui a changé depuis 25 ans ? », *Politiques et management public*, vol. 26/3, 2008.

Derouet Jean-Louis, « Entre la récupération des savoirs critiques et la construction des standards du management libéral. Recherche, administration et politique en France de 1975 à 2005 », *Revue des sciences de l'éducation*, n°32 (1), 2006.

Donegani Jean-Marie et Sadoun Marc, « La réforme de l'enseignement secondaire en France depuis 1945 : analyse d'une non-décision », *Revue française de science politique*, vol.26, n°6, 1976.

Dosière René, « Le contrôle ordinaire », *Pouvoirs*, vol 134, n°3, 2010.

Ducrot Olivier, « Illocutoire et performatif », *Dire et ne pas dire*, Paris, Hermann 3e édition : 1991.

Dupuy Claire, « Les politiques régionales d'éducation en Allemagne depuis PISA », *Revue internationale d'éducation de Sèvres*, n° 64, 2013.

Duran Patrice, « L'évaluation des politiques publiques : une irrésistible obligation », *Revue française des affaires sociales*, n°1-2, 2010.

Duran Patrice et Monnier Eric, « Le développement de l'évaluation en France. Nécessités techniques et exigences politiques », *Revue française de science politique*, n°2, 1992.

Duru-Bellat Marie, Mons Nathalie et Suchaut Bruno, « Organisation scolaire et inégalités sociales de performance : les enseignements de l'enquête PISA », *Education & formations* n°70, décembre 2004.

Dutercq Yves, Cuculou^[L]_[SEP]in Sylvie, Maroy Christian, *L'école à l'épreuve de la performance*, De Boeck Supérieur, 2013.

Duval René, « Quel cognitif retenir en didactique ? », *Recherches en Didactiques des Mathématiques*, 16(3), La Pensée Sauvage éditions, Grenoble, 1996.

Eggrickx Ariel, « Réflexion critique sur l'adoption d'outils de gestion par mimétisme : le cas de la LOLF », *Management et Avenir*, Management Prospective Ed., 4, n°54, 2012.

Emile-Besse Laurence, « Le système éducatif américain », *Revue internationale d'éducation de Sèvres*, n°35, avril 2004.

Emin Jean-Claude, « Réflexions sur le rôle de la Direction de l'évaluation et de la prospective (DEP) en France », *Education et sociétés*, 2, n°30, 2012.

Finck François, « L'indépendance statutaire des agents du conseil de l'Europe », *Revue française d'administration publique*, 2, n° 126, 2008.

Fouquet Annie, « L'évaluation des politiques publiques : État(s) de l'art et controverses », *Revue française d'administration publique*, vol. 4 n° 148, 2013.

Fumel Sylvie, Keskaik Saskia, « La motivation des élèves à répondre à un test standardisé : Résultats d'une étude dans le cadre de Cedre compétences langagières et littératie. *Éducation & formations*, DEPP, 2017.

Ganzeboom Harry B.G et al., « A Standard International Socio-Economic Index of Occupational Status », *Social Science Research*, 21, 1992.

Garraud Philippe, « Politiques nationales : l'élaboration de l'agenda », *L'Année sociologique*, 1990.

Gayon Vincent, « Le crédit vacillant de l'expert », *Cultures & Conflits*, n°75, automne 2009.

Gayon Vincent, « Un atelier d'écriture internationale : l'OCDE au travail. Éléments de sociologie de la forme « rapport » », *Sociologie du travail*, Vol. 51, n° 3, 2009.

Givord Pauline, Guillerm Marine, Monso Olivier, Murat Fabrice, « La ségrégation sociale entre les collèges. Quelles différences entre public et privé aux niveaux national, académique et local ? », *Education & formations* n°91, septembre 2016.

Granjou, Céline. « L'expertise scientifique à destination politique », *Cahiers internationaux de sociologie*, vol. 114, n°1, 2003.

Grapin Nadine et Sayac Nathalie, « Evaluer la maîtrise de la numération écrite chiffrée : choix du CQM et validité d'items d'évaluations externes », *Education et didactique*, 11-3, 2017.

Grapin Nadine et Sayac Nathalie, « Evaluer les capacités des élèves à résoudre des problèmes dans le cas d'une évaluation externe en France : les spécificités de la forme QCM », *Education et francophonie*, XLII(2), 2014.

Grapin Nadine et Sayac Nathalie, « Evaluer par QCM en fin d'école : stratégies et degrés de certitude », *Annales de didactique et de sciences cognitives*, 19, 2014.

Groux Dominique, « L'éducation comparée : approches actuelles et perspectives de développement », *Revue française de pédagogie*, volume 121, 1997.

Grugeon-Allys Brigitte, Grapin Nadine, « Validité d'une évaluation externe. Complémentarité des approches didactique et psychométrique ». *Actes du séminaire national de l'Association pour la Recherche en Didactique des Mathématiques*, mars 2015.

Haas Peter M., « Epistemic Communities and International Policy Coordination », *International Organization*, 46/1, 1992.

Hall Peter, « Policy Paradigms, Social Learning, and the State : The Case of Economic Policymaking in Britain », *Comparative Politics*, Vol.25, n°3, 1993.

Halpern Charlotte et Le Galès Patrick, « Pas d'action publique autonome sans instruments propres. Analyse comparée et longitudinale des politiques environnementales et urbaines de l'Union européenne », *Revue Française de Science Politique*, 61 (1), 2011.

Henry Odile, Pierru Frédéric, « Les consultants et la réforme des services publics », *Actes de la recherche en sciences sociales*, n°193, 2012.

Houillon Philippe, « Le contrôle extraordinaire du parlement », *Pouvoirs*, 3, n°134, 2010.

Huteau Michel, « Alfred Binet et la psychologie de l'intelligence », *Le journal des psychologues*, 1, n°234, 2006.

Jobard Fabien, et al. « Sociologie politique des passeurs. Acteurs dans la circulation des savoirs, des normes et des politiques publiques », *Revue française de science politique*, vol. 70, n° 5, 2020.

Julo Jean « Des apprentissages spécifiques pour la résolution de problèmes ? », *Grand N*, 69, 2002.

Julo Jean, *Représentation des problèmes et réussite en mathématiques*, Rennes, Presses universitaires de Rennes, 1995.

Keskpaik Saskia et Rocher Thierry, « La mesure de l'équité dans PISA : pour une décomposition des indices statistiques », *Education & Formations* n°80, décembre 2011.

Keskpaik Saskia et Rocher Thierry, « La motivation des élèves français face à des évaluations à faibles enjeux. Comment la mesurer ? Son impact sur les réponses. » *Education & Formations*, n°86-87, mai 2015.

Keslair François, « Analysing PIAAC Data with Stata », dans Debora Maehler, Beatrice Rammstedt (dir.), *Large-Scale Cognitive Assessment. Methodology of educational measurement and assessment*, Springer Cham, 28 juillet 2020.

Kieffer Annick et Tréhin-Lalanne Rémi, « La fabrication d'un consensus : la révision de la Classification Internationale Type de l'Education », *Sociologie et sociétés*, Vol.43, 2, automne 2011.

Kott Sandrine, « Une "communauté épistémique" du social ? Experts de l'OIT et internationalisation des politiques sociales dans l'entre-deux guerres », *Genèses*, 2/71, 2008.

Kott Sandrine, « Dynamiques de l'internationalisation : l'Allemagne et l'Organisation internationale du travail (1919-1940) », *Critique internationale*, n°52, 2011.

Lacouette Fougère Clément et Lascoumes Pierre, « L'évaluation : un marronnier de l'action gouvernementale ? », *Revue française d'administration publique*, vol. 148, n° 4, 2013.

Laderriere Pierre, « Organismes d'enseignement et de recherche », *Revue française de pédagogie*, volume 50, 1980.

Lapostolle Guy, « Du bon usage de l'expertise dans les politiques éducatives », *Politiques et management public*, Vol. 24, n°1, 2006.

Lascombe Michel et Xavier Vandendriessche. « La loi organique relative aux lois de finances (LOLF) et le contrôle des finances publiques », *Revue française d'administration publique*, vol.117, n°1, 2006.

Lascoumes Pierre, « L'expertise, de la recherche d'une action rationnelle à la démocratisation des connaissances et des choix », *Revue française d'administration publique*, 3, n°103, 2002.

Lauwerier Thibaut, « L'éducation au service du développement. La vision de la Banque mondiale, de l'OCDE et de l'UNESCO. » *L'éducation en débats : analyse comparée*, vol.9, 2017.

Le Donné Noémie, Rocher Thierry, « Une meilleure mesure du contexte socio-éducatif des élèves et des écoles – Construction d'un indice de position sociale à partir des professions des parents », *Education & Formations* n°79, décembre 2010.

Leclercq Véronique, « La notion de littéracie dans le champ des formations de base pour adultes : référence explicite ou influence diffuse ? », *Spirale, revue de recherches en éducation*, n°53, 2014.

Lefresne Florence, « Du bon usage des comparaisons internationales dans l'aide au pilotage des systèmes éducatifs », *Education & Formations*, DEPP, 2016.

Léger-Jarniou Catherine, « Développer la culture entrepreneuriale chez les jeunes », *Revue française de gestion*, 5, n°185, 2008.

Legrand André, « L'éducation nationale, une institution rétive à la réforme ? », *Politiques et management public*, 2005.

Lingard Bob, « Le « panoptisme global » : l'influence des tests internationaux et des infrastructures de données en éducation », *Administration & Éducation*, vol. 145, n°1, 2015.

Lingard Bob, Martino Wayne et Rezai-Rashti Goli, « Testing Regimes, Accountabilities and Education Policy : Commensurate Global and National Developments », *Journal of Education Policy*, vol.28, 5, 2013.

Mangez Eric et Cattonar Branka, « A quoi servent les enquêtes PISA ? Réflexions sur le cas de la Belgique francophone », *Sciences de la société*, n°79, 2011.

Mankiw G., Romer D. et Weil D., « A Contribution to the Empirics of Economic Growth », *Quarterly Journal of Economics*, 1992.

Margaria Claire, Simonis-Sueur Caroline, « La revue fête son centième numéro! : regard rétrospectif sur près de 40 ans de publication. », *Education & formations*, DEPP, 2019.

Maroy Christian et Mangez Catherine, « Rationalisation de l'action publique ou politisation de la connaissance ? », *Revue française de pédagogie*, 164, 2008.

Matasci Damiano, « L'éducation, terrain d'action internationale : le Bureau international de l'enseignement technique dans les années 1930 », *Relations internationales*, 3, n° 151, 2012.

McGaw Barry, «The role of the OECD in international comparative studies of achievement », *Assessment in Education: Principles, Policy & Practice*, 15:3, 2008.

Merle Pierre, « L'efficacité de l'enseignement », *Revue française de sociologie*, vol. 39, n° 3, 1998.

Meyer Morgan et Molyneux-Hodgson Susan, « « Communautés épistémiques » : une notion utile pour théoriser les collectifs en sciences ? », *Terrains & travaux*, vol. 18, n°1, 2011.

Mezeix Jean-François et Grange Catherine, « Les facteurs influençant l'orientation et le parcours de la troisième au post-baccalauréat d'une cohorte d'élèves de l'académie de Clermont-Ferrand », *Education & formations* n°77, novembre 2008.

Michel Alain et Mons Nathalie, « PISA et les politiques éducatives : une influence complexe », *Administration & Education*, vol.145, n°1, 2015.

Michel Alain et Pons Xavier, « Éditorial. Le « choc » PISA : métaphore ou révolution ? », *Administration & Éducation*, vol.145, n°1, 2015.

Mons Nathalie et Pons Xavier, « Pourquoi n'y a-t-il pas de choc PISA en France ? Sociologie de la réception d'une enquête internationale dans trois pays européens (2001-2008) », *Revue française de pédagogie*, n°182, janvier-mars 2013.

Mons Nathalie, « L'évaluation des politiques éducatives. Apports, limites et nécessaire renouvellement des enquêtes internationales sur les acquis des élèves », *Revue internationale de politique comparée*, 14, 3, 2007.

Mons Nathalie, « Evaluation des politiques éducatives et comparaisons internationales », *Revue française de pédagogie*, n°164, juillet septembre 2008.

Mons Nathalie, « Effets théoriques et réels des politiques d'évaluation standardisées », *Revue française de pédagogie*, 169, octobre-novembre-décembre, 2009.

Morgan Clara, « Tracing the Sub-national Effect of the OECD PISA : Integration into Canada's Decentralized Education System », *Global Social Policy*, 16 :1, 2015.

Muller Pierre-Eugène, « De l'instruction publique à l'éducation nationale », dans *Mots*, n°61, décembre 1999.

Muller Pierre-Eugène, « L'analyse cognitive des politiques publiques : vers une sociologie politique de l'action publique », *Revue française de science politique*, 50^e année, n°2, 2000.

Nelson Phillip, « Information and Consumer Behaviour », *Journal of Political Economy*, vol. 78, 1970.

Neveu Erik, « L'analyse des problèmes publics. Un champ d'étude interdisciplinaire au cœur des enjeux sociaux présents », *Idées économiques et sociales*, vol. 190, n°4, 2017.

Neveu Erik, « La construction sociale des problèmes publics », *Idées économiques et sociales*, n°190, décembre 2017.

Niemann Dennis et Martens Kerstin, « When Do Numbers Count ? The Differential Impact of Ratings and Rankings on National Education Policy in Germany and the US », *German Politics*, 22 (3), 2013.

Niemann Dennis, « Le « choc PISA » en Allemagne et les réformes de l'éducation », *Administration & Éducation*, vol. 145, n°1, 2015.

Normand Romuald, Vincent-Dalud Monique, « Sciences de gouvernement de l'éducation et réseaux transnationaux d'experts : la fabrication d'une politique européenne », *Éducation et société*, 29 (1), 2012.

Novoa Antonio et Yariv-Mashal Tali, « Comparative Research in Education : a mode of governance or a historical journey ? » *Comparative Education*, vol. 39, 4, 2003.

Pénissat Etienne, « Publier des « chiffres officiels » ou les contraintes bureaucratiques et politiques qui façonnent l'expertise d'État. Le cas des statistiques du ministère du Travail », *Gouvernement et action publique*, Vol. 1, n°4, 2012

Perret Bernard, « De l'échec de la rationalisation des choix budgétaires (RCB) à la loi organique relative aux lois de finances (LOLF) », *Revue française d'administration publique*, 1, n°117, 2006.

Pisani-Ferry Jean, Sapir André, « Last Exist to Lisbon », *Bruegel Policy Contribution*, mars 2006.

Pollitt Christopher, « Performance information for democracy : The missing link ? » *Evaluation*, 12 (1), 2006.

Pons Xavier, « Qu'apprend-on vraiment de PISA ? Sociologie de la réception d'une enquête internationale dans trois pays européens (2001-2008) », *Revue internationale d'éducation de Sèvres*, 54, 2010.

Pons Xavier, « Figures de l'expert international français : l'exemple des évaluations internationales des acquis (1958-2008) », « Carrefours de l'éducation », n°37, Armand Colin, 2014.

Pons Xavier, « Les statisticiens du ministère de l'Éducation nationale : évolution d'un métier d'État 1957-2007) », *Histoire de l'éducation*, 140-141, 2014.

Pons Xavier, « PISA au Parlement ou l'effet boomerang d'un outil de communication politique », *Administration & Éducation*, vol. 145, n°1, 2015.

Rakocevic Robert, « Implication des parents dans la réussite à l'école : éclairages internationaux », *Education & formations* n°85, novembre 2014.

Raulin Dominique, « De nouveaux rapports entre science et politique : le cas des programmes scolaires », *Revue française de pédagogie*, 154, 2006.

Ravinet Pauline, « Fenêtre d'opportunité », Laurie Boussaguet éd., *Dictionnaire des politiques publiques. 3e édition actualisée et augmentée*, Presses de Sciences Po, 2010.

Restier-Melleray Christiane, « Experts et expertise scientifique. Le cas de la France », *Revue française de science politique*, 40-4, 1990.

Rocher Thierry et Cosnefroy Olivier, « Le redoublement au cours de la scolarité obligatoire : nouvelles analyses, mêmes constats », *Education & formation* n°70, décembre 2004.

Rocher Thierry, « Comparaisons internationales », *Education & Formation* n°78 numéro spécial, novembre 2008.

Rocher Thierry, « Que nous apprennent les évaluations internationales sur le fonctionnement des systèmes éducatifs ? Une illustration avec la question du redoublement », *Education & Formations* n°78, novembre 2008.

Rocher Thierry, « Mesure des compétences. Méthodes psychométriques utilisées dans le cadre des évaluations des élèves », *Education & Formations*, n°86-87, mai 2015.

Rocher Thierry, « PISA, une belle enquête : lire attentivement la notice ». Association Française des Acteurs de l'Éducation, *Administration & Education*, vol.145, n°1, 2015.

Romainville Marc, « Du bon usage de PISA », *La revue nouvelle*, vol. 115, issue 3-4, 2002.

Roqueplo Philippe, « L'expertise scientifique, consensus ou conflit ? », *La terre outragée, les experts sont formels*, Autrement, série sciences en société, n°1, 1992.

Rubner Jeanne, « Les suites du choc PISA dans les médias allemands », *Revue internationale d'éducation de Sèvres*, n°66, 2014.

Rutgers Mark et Mentzel Maarten A., « Scientific expertise and public policy : resolving paradoxes ? », *Science and Public Policy*, 26(3), 1999.

Santos Iris et Kauko Jaakko, « Externalisations in the Portuguese parliament : analysing power struggles and (de-)legitimation with Multiple Streams Approach », *Journal of Education Policy*, juin 2020.

Sauvageot Claude, « Comparaisons internationales », *Education & Formation* n°78 numéro spécial, novembre 2008.

Sawicki Frédéric, « Pour une sociologie des problématisations politiques de l'école », *Politix*, n°98, 2012.

Simola Hannu, Varjo Janne, Rinne Risto, « A contre-courant : dépendance au sentier, convergence et contingence. Vers une meilleure compréhension du modèle finlandais d'assurance-qualité et d'évaluation », *Education et sociétés*, 2 n°28, 2011.

Speer Sandra, Pattyn Valérie, De Peuter Bart, « L'importance grandissante des évaluations au sein des parlements : Les gouvernements tenus de rendre des comptes ? » *Revue Internationale des Sciences Administratives*, vol. 81, n°1, 2015.

Steinert Brigitte, « Evaluation et système-qualité, Panorama en Allemagne », *Revue internationale d'éducation de Sèvres*, juin 2000.

Thélot Claude, « L'évaluation du système éducatif français », *Revue française de pédagogie*, volume 107, 1994.

Thiers Eric, « Le contrôle parlementaire et ses limites juridiques : un pouvoir presque sans entraves », *Pouvoirs*, vol 134, n°3, 2010.

Tiana-Ferrer Alejandro, « Pilotage par les résultats et amélioration de l'éducation : deux réalités indissociables ? », *Administration et éducation*, n°98, 2003.

Tillmann Klaus-Jürgen, Dederich Kathrin, Kneuer Daniel, Kullmann Christian et Nessel Isa, « PISA als bildungspolitisches Ereignis. Fallstudien in vier Bundesländern », *Wiesbaden : VS Verlag für Sozialwissenschaften*; 2008.

Tournier Michèle, « La formation des enseignants par la recherche, en France : quelques points de repère (1966-1986) », dans *Recherche & Formation*, n°2, 1987.

Vallet Louis-André, « L'enquête Information et vie quotidienne 2011 parmi les dispositifs nationaux et internationaux de mesure des compétences des adultes », *Economie et statistique*, n°490, 2016.

Vrignaud Pierre, « La mesure de la littératie dans PISA : la méthodologie est la réponse, mais quelle était la question ? », *Education & Formations* n°78, novembre 2008.

Articles de presse et communiqués

Fohr Anne, « Le thermomètre de l'école est bloqué », *Nouvel observateur*, avril 1999.

Interview de M. François Fillon, ministre de l'éducation nationale, de l'enseignement supérieur et de la recherche, dans "Métro" du 6 décembre 2004.

Le Monde, « Une concertation sur l'école très encadrée », 5 juillet 2012

Le Point, « Peillon et Ayrault lancent la concertation pour « refonder l'école », 5 juillet 2012.

SNUipp, « 100 minutes sans convaincre », *Communiqué de presse* du 19 novembre 2004.

Tract commun CGT, CFDT et SNUTETE FSU de l'administration centrale du Ministère du travail, de l'emploi et de la santé (DARES et DREES), CGT Educ'action centrale du Ministère de l'Education nationale et Ministère de l'enseignement supérieur et recherche, SYGMA FSU du Ministère de l'Agriculture et de la pêche (SCEES), CGT Equipement-environnement de l'Administration Centrale du Ministère de l'Ecologie, du Développement durable, des Transports et du Logement, CGT, CFDT, SUD de l'Insee, CGT de Pôle Emploi, *L'indépendance de la DEPP (Direction de l'Evaluation, de la Prospective et de la Performance) mise à mal : le débat faussé sur l'Education nationale*, publié le 2 novembre 2011.

Weisner François, *La LOLF*, document publié par le syndicat SUD éducation le 14 décembre 2004.

Rapports et manuels

Allain Joël et Devilliers Michel, *Rapport sur la situation des publications de la DEPP*, INSEE, 6 février 2012.

Andreu Sandra, Eteve Yann, Garcia Emilie, Rocher Thierry et Vourc'h Ronan, *Rapport technique CEDRE*, sciences expérimentales, DEPP, 2013.

Bottani Norbert et Vrignaud Pierre, *La France et les évaluations internationales*, Rapport national au HCEE, Paris, 2005.

Bourdieu Pierre et Gros François, *Principes pour une réflexion sur les contenus de l'enseignement*. Commission présidée par Pierre Bourdieu et François Gros, Paris, Imprimerie Nationale, 1989.

Bourny Ginette, Braxmeyer Nicole, Dupé Claire, Rémond Martine, Robin Isabelle et Rocher Thierry, *Les compétences des élèves français à l'épreuve d'une évaluation internationale, premiers résultats de l'enquête PISA 2000*, DPD, 2001.

Comité de pilotage de la concertation sur la refondation de l'école, *Refondons l'école de la république*, 2012.

Commission du débat national sur l'avenir de l'école, *Pour la Réussite de tous les élèves*, La Documentation Française, Paris, 2004.

Commission du débat national sur l'avenir de l'école, *Les Français et leur école. Le miroir du débat*, Dunod, 2004.

Commission nationale du débat public, *Ecole et collège : tout ce que nos enfants doivent savoir*, CNDP/XO éditions, 2006.

Conseil scientifique de l'éducation nationale, *Présentation et activités 2019-2020*, Ministère de l'éducation nationale, 2020.

CREDOC, *Performance, efficacité, efficience : les critères d'évaluation des politiques sociales sont-ils pertinents ?* Cahier de recherche, n°299, décembre 2012.

Cour des Comptes, *La gestion des enseignants du premier degré*, 1990.

Cour des Comptes, *L'organisation du baccalauréat et les IUT*, 1994.

Cour des Comptes, *Les effets de la décentralisation dans l'enseignement secondaire*, 1995.

Cour des Comptes, *La gestion du système éducatif*, avril 2003.

Cour des Comptes, *L'Education nationale : organiser son évaluation pour améliorer sa performance*, décembre 2017.

Delahaye Jean-Paul, Mauhourat Marie-Blanche, Monnanteuil François, Poncelet Yves, Vibert Anne, Yebbou Johan, Inspection générale de l'éducation

nationale, *Les évaluations nationales et internationales sur les acquis des élèves et sur d'autres dimensions des systèmes éducatifs*, rapport n°2012-056 à Monsieur le ministre de l'éducation nationale, mai 2012.

DEPP, Numéro spécial : 18 questions sur le système éducatif : synthèses de travaux de la DEP, *Education & formations* n°66 juillet-décembre 2003.

DEPP, « La maîtrise du langage et de la langue française à la fin de l'école primaire », *Note d'évaluation* n° 04.10, octobre 2004.

DEPP, « L'évaluation internationale PISA 2003 : compétences des élèves français en mathématiques, compréhension de l'écrit et sciences », *les Dossiers Enseignement scolaire*, n°180, mars 2007.

DEPP, *L'état de l'école : coûts ; activités ; résultats : 31 indicateurs sur le système éducatif français*, n°22, Ministère de l'Éducation nationale, octobre 2012.

DEPP, *Cycle des évaluations disciplinaires réalisées sur échantillons, rapport technique, Mathématiques 2014*, Ecole, décembre 2015.

DEPP, *Evaluation numérique des compétences du socle en début de sixième : des niveaux de performance contrastés selon les académies*, Note d'information n°18, 2016.

Dulot Alain, *Refondons l'école de la république*. Rapport de la concertation, 2013.

Echazarra Alfonso, « Les jeunes de 15 ans sont-ils devenus plus « verts » ? », *PISA à la loupe* n°87, OCDE, 28 novembre 2018.

Givord Pauline, « Le renforcement de la mixité sociale à l'école a-t-il un impact sur l'équité des résultats d'apprentissage ? », *PISA à la loupe* n°97, OCDE, 11 juin 2019.

Givord Pauline, « Quels liens entre les politiques de libre choix des établissements et la mixité sociale à l'école ? », *PISA à la loupe* n°96, OCDE, 15 mai 2019.

Grek Sotiria, Lawn Martin et Ozga Jenny, *Study on the Use and Circulation of PISA in Scotland*, Rapport pour le projet Know&Pol, Université d'Edinburgh, Centre d'études sociologiques, 2009.

Haut conseil de l'éducation, *Recommandations pour le socle commun*, 23 mars 2006.

Haut conseil de l'éducation, *Mise en œuvre du socle commun. Bilan des résultats de l'Ecole*, 2011.

Haut conseil de l'évaluation de l'école, *Eléments pour un diagnostic sur l'Ecole*. Document général préparatoire au débat national sur l'avenir de l'école, 2003.

Haut conseil de l'évaluation de l'école, *Le redoublement permet-il de résoudre les difficultés rencontrées au cours de la scolarité obligatoire ?*, avis n°14, décembre 2004 ; à partir du rapport établi par Jean-Jacques Paul et Thierry Troncin.

Haut conseil de l'évaluation de l'école, *La France et les évaluations internationales des acquis des élèves*, avis n° 16, mai 2005.

Haut conseil de l'évaluation de l'école, *Les dispositifs d'indicateurs internationaux de l'éducation*, avis n°18, juillet 2005.

INSEE, *Portrait social*, édition 2016.

Inspection générale de l'éducation nationale, *Les évaluations nationales et internationales sur les acquis des élèves et sur d'autres dimensions des systèmes éducatifs*, rapport n°2012-056, mai 2012.

Institut national de recherche pédagogique, *Standards, compétences de base et socle commun*, décembre 2005, rédacteur Olivier Meunier.

Ministère de l'Éducation nationale, *Loi d'orientation sur l'École. 14 propositions de François Fillon*, novembre 2004.

Mission sur l'Évaluation et l'organisation de la fonction statistique au Ministère de l'Éducation nationale, de l'enseignement supérieur et de la recherche (MENESR), *Évaluation et organisation de la fonction statistique au ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche*, mai 2016.

Mons Nathalie et Pons Xavier, avec la participation d'Agnès Van Zanten et Jeremy Pouille, *La réception de PISA en France, Connaissances et régulation du système éducatif*, KNOWandPOL, orientation 3, avril 2009.

Mostafa Tarek, « Pourquoi les filles ne sont-elles pas plus nombreuses à choisir une carrière scientifique ? », *PISA à la loupe* n°93, OCDE, 17 septembre 2019.

Murat Fabrice et Rocher Thierry, « La place du projet professionnel dans les inégalités de réussite scolaire à 15 ans », *France, portrait social 2002-2003*, INSEE ;

OCDE, *Regards sur l'Éducation*, Editions OCDE, 1993.

OCDE, publications des résultats des cycles PISA 2000, 2003, 2006, 2009, 2012, 2015, 2018 et de leurs annexes, Editions OCDE.

OCDE, *Analyse des politiques d'éducation*. Editions OCDE, 1997.

OCDE, *Résolution du conseil instituant un programme d'élaboration périodique d'indicateurs sur les acquis des élèves*, 15 décembre 1997.

OCDE-PISA, *Consignes pour la traduction du matériel PISA*. Document provisoire préparé par Aletta Grisay, octobre 1998.

OCDE-UNESCO-UIS, *Literacy Skills for the World of Tomorrow, further results from PISA 2000*, 2003.

OCDE, *Analyse des politiques d'éducation*. Editions OCDE, 2005.

OCDE, *Définition et Sélection des Compétences clés*, Editions OCDE, 2005.

OCDE, *Elaborer une stratégie en faveur des compétences*, réunion du Conseil au niveau des ministres, Paris, 25 et 26 mai 2011.

OCDE, *L'éducation aujourd'hui. La perspective de l'OCDE*, Editions OCDE, 2013.

OCDE, *Politiques meilleures, France*, Editions OCDE, 2015.

OCDE, *PISA National Project Manager*, Editions OCDE, 10 mars 2017.

OCDE, *Statut, règlement et instructions applicables aux agents de l'organisation*, actualisation de mars 2020.

OCDE, *Perspectives des politiques de l'éducation – France*, Editions OCDE, juin 2020.

Pair Claude, *Les forces et les faiblesses de l'évaluation du système éducatif français*, HCée, 2001.

Parlement européen, *Recommandation du Conseil européen et du Parlement européen*, 2006.

Pons Xavier, *The French Contribution to the International Design of Pisa*, Document de travail pour le projet Know&Pol, Paris : OSC, février 2009.

Rapport d'information n°1275 *sur la question du port des signes religieux à l'école* fait par M. Jean-Louis Debré, député, Assemblée Nationale, 4 décembre 2003.

Rapport n°2025 fait au nom de la commission des affaires culturelles, familiales et sociales *sur le projet de loi d'orientation pour l'avenir de l'école*, par M. Frédéric Reiss, député, enregistré à la Présidence de l'Assemblée nationale le 9 février 2005.

Rapport n°234 fait au nom de la commission des affaires culturelles, *sur le projet de loi d'orientation pour l'avenir de l'école*, par M. Jean-Claude Carle, sénateur, 9 mars 2005.

Rapport d'information n°4364 de MM. Philippe Cochet et Marc Dolez *sur l'évaluation des incidences de la Stratégie de Lisbonne sur l'économie française*, comité d'évaluation et de contrôle des politiques publiques, 16 février 2012.

Rapport n°767 fait au nom de la commission des affaires culturelles et de l'éducation *sur le projet de loi d'orientation et de programmation pour la refondation de l'école de la République*, Tome 1, par Yves Durand, député, enregistré à la Présidence de l'Assemblée nationale le 28 février 2013.

Rapport n°568 fait au nom de la commission de la culture, de l'éducation et de la communication *sur le projet de loi d'orientation et de programmation pour la refondation de l'école de la République*, par Françoise Cartron, sénatrice, enregistré à la Présidence du Sénat le 14 mai 2013.

Rapport fait au nom de la commission d'enquête *sur l'exil des forces vives de France*, présidé par Luc Chatel, Rapporteur Yann Galut, 8 octobre 2014.

Rapport d'information n°2535 du comité d'évaluation et de contrôle des politiques publiques *sur l'évaluation des politiques publiques en faveur de la mobilité sociale des jeunes*, déposé le 29 janvier 2015 la Présidence de l'Assemblée nationale et présenté par Régis Juanico et Jean Frédéric Poisson, députés.

Rapport d'information n°2620 déposé par la commission des affaires européennes le 4 mars 2015 *sur l'emploi des jeunes en Europe*, présenté par Philip Cordery, député.

Rapport d'information n°590 fait au nom de la commission d'enquête *sur le fonctionnement du service public de l'éducation, sur la perte de repères républicains que révèle la vie dans les établissements scolaires et sur les difficultés rencontrées par les enseignants dans l'exercice de leur profession*, par Jacques Gersperrin, sénateur, enregistré à la présidence du Sénat le 1^{er} juillet 2015.

Rapport d'information n°2951, déposé par la commission des affaires culturelles et de l'éducation en conclusion des travaux de la mission *sur les liens entre le lycée et l'enseignement supérieur*, présenté par Emeric Bréhier, député, le 8 juillet 2015.

Rapport d'information n°3292 du comité d'évaluation et de contrôle des politiques publiques *sur l'évaluation des politiques publiques en faveur de la mixité sociale dans l'éducation nationale*, déposé le 1^{er} décembre 2015 à la Présidence de l'Assemblée nationale et présenté par Yves Durand et Rudy Salles, députés.

Rapport d'information n°1265 du comité d'évaluation et de contrôle des politiques publiques de l'Assemblée nationale, *sur l'organisation de la fonction d'évaluation du système éducatif*, déposé le 27 septembre 2018 à la Présidence de l'Assemblée nationale et présenté par Régis Juanico et Marie Tamarelle-Verhaeghe, députés.

Robin Isabelle et Rocher Thierry, « La compétence en lecture des jeunes de 15 ans : une comparaison internationale », *Données sociales : la société française – Edition 2002-2003*, INSEE.

Rocher Thierry, Fumel Sylvie, Chesné Jean-François, « Méthodologie de l'évaluation des compétences de base en français et en mathématiques en fin d'école et en fin de collège », *Note d'information n°37*, DEPP, décembre 2008.

Lois, textes règlementaires, débats parlementaires, lettres de mission, communications institutionnelles

Lois

Annexe à la loi n°2013-595 du 8 juillet 2013 d'orientation et de programmation pour la refondation de l'école de la République.

Annexe budgétaire au Projet de loi de finance 2019, mission Enseignement scolaire, Programme 214, Soutien de la politique de l'éducation nationale, Action n°02 Evaluation et contrôle.

Loi n°75-620 du 11 juillet 1975 relative à l'éducation.

Loi n°83-634 du 13 juillet 1983 portant droits et obligations des fonctionnaires.

Loi n° 2005-380 du 23 avril 2005 d'orientation et de programme pour l'avenir de l'École.

Loi n°2013-595 du 8 juillet 2013 d'orientation et de programmation pour la refondation de l'école de la République.

Projet de loi d'orientation pour l'avenir de l'école n°2025, enregistré à la Présidence de l'Assemblée nationale le 12 janvier 2005.

Projet de Loi de Finance 2006, Bleu budgétaire de la mission Enseignement scolaire. Programme 214 : Soutien de la politique de l'éducation nationale

Projet de Loi de Finance 2006, Bleu budgétaire de la mission Immigration, asile et intégration, Programme 104 : Intégration et accès à la nationalité française.

Projet de Loi de Finance 2010, Bleu budgétaire de la mission Enseignement scolaire. Programme 214 : Soutien à la politique du Ministère de l'éducation nationale, justification au premier euro.

Textes réglementaires et conventionnels

Annexe intégrée à la section A du chapitre II du titre II du livre 1er du code de l'éducation, par le décret n°2006-830 du 11 juillet 2006, signé par Dominique de Villepin, Premier ministre, Gilles de Robien, ministre de l'éducation nationale, de l'enseignement supérieur et de la recherche, et François Baroin ministre de l'outre-mer.

Circulaire ministérielle du 23 février 1989 relative au renouveau du service public, JORF, 24 février 1989.

Convention relative à l'Organisation de Coopération et de Développement Economiques, faite à Paris le 14 décembre 1960. Décret du 10 février 1987.

Décret n°87-83 du 10 février 1987 portant organisation de l'administration centrale du ministère de l'Education nationale.

Décret n°2005-999 du 22 août 2005 relatif au Haut Conseil de l'éducation mettant en œuvre la loi d'orientation.

Décret n° 2014-133 du 17 février 2014 fixant l'organisation de l'administration centrale des ministères de l'éducation nationale et de l'enseignement supérieur et de la recherche.

Règlement CE n°223/2009 du Parlement européen et du Conseil européen du 11 mars 2009 relatif aux statistiques européennes, modifié par le règlement UE 2015/759 du Parlement européen et du Conseil européen du 29 avril 2015.

Débats, interventions, auditions parlementaires

Compte-rendu de l'audition de François Fillon à l'Assemblée nationale le 8 février 2005, JORF.

Compte-rendu de l'audition publique de Luc Ferry et Xavier Darcos devant la commission des affaires culturelles du Sénat, 3 juillet 2003, JORF.

Comptes-rendus des débats en séance, Sénat, séance du 27 juin 1989, JORF.

Comptes-rendus des débats en séance, Assemblée nationale, 2e séance du 14 février 2005, JORF.

Comptes-rendus des débats en séance, Assemblée nationale, 2 séance du 15 février 2005, JORF.

Comptes-rendus des débats en séance, Assemblée nationale, 3e séance, 15 février 2005, JORF.

Comptes-rendus des débats en séance, Assemblée nationale, 1ere séance du 17 février 2005, JORF.

Comptes-rendus des débats en séance, Assemblée nationale, 1ere séance du 11 mars 2013, JORF.

Comptes-rendus des débats en séance, Assemblée nationale, 1ere séance du 12 mars 2013, JORF.

Comptes-rendus des débats en séance, Assemblée nationale, 2e séance du 13 mars 2013, JORF.

Comptes-rendus des débats en séance, Sénat, séance du 21 mai 2013, JORF.

Comptes-rendus des débats en séance, Sénat, séance du 22 mai 2013, JORF.

Conseil des ministres, exposé des motifs du projet de loi pour l'avenir de l'école, présenté au nom de M. Jean-Pierre Raffarin, Premier ministre, par M. François Fillon, Ministre de l'éducation nationale, de l'enseignement supérieur et de la recherche, le 12 janvier 2005.

Question orale n°0692S de Mme Nicole Borvo (Paris, CRC) publiée dans le JO du Sénat du 19 janvier 2000.

Question orale n°613 de Madame Cécile Gallez, Nord, UMP, publiée au JO de l'Assemblée Nationale le 14 mai 2003.

Question écrite n°08098 de Madame Danièle Pourtant, Paris, Socialiste, publiée au JO du Sénat le 19 juin 2003 et réponse du Ministère de la jeunesse, de l'éducation nationale et de la recherche publiée au JO du Sénat le 7 août 2003.

Question écrite n°56726 de Monsieur Christian Estrosi posée à Monsieur le ministre de l'Education nationale publiée au JO de l'Assemblée Nationale le 1er janvier 2005.

Question écrite n°93873 posée par Denis Merville, UMP, publiée au journal officiel de l'Assemblée Nationale le 9 mai 2006.

Question cible thématique (question orale) n°0156C de Madame Jacqueline Gourault, publiée au JO du Sénat le 9 février 2011.

Question écrite n°03226 de Monsieur Gaëtan Gorce, Nièvre, Socialiste, publiée au JO du Sénat le 22 novembre 2012.

Question écrite n°09799 de M. Robert Tropeano, Hérault, RDSE, publiée au JO du Sénat le 19 décembre 2013.

Démocrates Indépendants, publiée au JO de l'Assemblée Nationale le 4 décembre 2013.

Question écrite n°48629 posée par la députée Marietta Karamanli, Sarthe, Socialiste, publiée au JO de l'Assemblée Nationale le 4 février 2014.

Question au gouvernement n°0504G de M. Jacques-Bernard Magner, Puy-de-Dôme, socialiste, publiée dans le JO du Sénat du 17 avril 2015.

Question n°2976 de Monsieur Jean-Louis Costes, Lot-et-Garonne, Les Républicains, publiée au JO de l'Assemblée nationale le 10 juin 2015.

Question au gouvernement n°1069G de Monsieur Gérard Longuet, Meuse, Les Républicains, publiée au JO du Sénat le 16 décembre 2016.

Question au gouvernement n°1109G de Monsieur Alain Marc, Aveyron, Les Républicains, publiée au JO du Sénat le 18 janvier 2017.

Question écrite n°6388 de M. Eric Pauget, Les Républicains, publiée au JO de l'Assemblée Nationale le 13 mars 2018.

Question au gouvernement n°1017G de la sénatrice Nathalie Delattre, Gironde, RDSE, publiée au JO le 7 novembre 2019.

Question écrite n°19870 de M. Philippe Paul, Finistère, Les Républicains, publiée au JO du Sénat le 31 décembre 2020.

Réponse du ministre de l'Éducation nationale à la question au gouvernement n°1528 posée par Monsieur Hervé Fréron, Meurthe-et-Moselle, groupe socialiste, radical, citoyen et divers gauche, publiée au JO de l'Assemblée nationale le 16 janvier 2014.

Réponse du ministre de l'éducation nationale à la question cible thématique n°0153C de Jacques Legendre, UMP, publiée au journal officiel du Sénat le 9 février 2011.

Réponse du ministre de l'Education nationale Luc Chatel à la question crible thématique (question orale) n°0155C posée par la sénatrice Brigitte Gonthier Maurin, Hauts-de-Seine, Parti Communiste Français, publiée au JO du Sénat le 9 février 2011.

Réponse du ministre Vincent Peillon à la question posée au gouvernement n°1402 par Vincent Feltesse, socialiste, publiée au journal officiel de l'Assemblée nationale le 14 décembre 2013.

Réponse du ministre à la question orale n°0974S de Madame Aline Archimbaud, Seine-Saint-Denis, Ecologiste, publiée au JO du Sénat le 25 décembre 2014.

Lettres de mission

Lettre de mission du Premier ministre à Patrick Viveret, chargé de mission auprès du Commissaire au Plan, portant commande de propositions en vue de l'évaluation du revenu minimum d'insertion, 22 juillet 1988.

Lettre de mission du 24 janvier 2014 de Vincent Peillon, ministre de l'Education nationale, à Nathalie Mons, Présidente du CNESCO.

Chartes

Charte de l'évaluation des politiques publiques et des programmes publics adoptée en juin 2006 par les membres de la Société Française de l'Evaluation

Communications institutionnelles

Conclusions du Conseil Européen de Lisbonne, 23 et 24 mars 2000.

Discours et interventions publiques

Abbé Grégoire, Discours sur l'éducation commune, 30 juillet 1793.

Gurría Angel, Discours du 24 mai 2011 de M. le Secrétaire général de l'OCDE, www.oecd.org/fr/apropos/secretairegeneral/betterpoliciesforbetterlives.htm

Hollande François, Discours de M. le Président de la République à la suite de la concertation sur l'Ecole, prononcé le 12 octobre 2012.

Monory René, ministre de l'Éducation nationale, *Bonne rentrée*, « cahiers de l'éducation nationale », septembre 1986.

Sarkozy Nicolas, Discours d'Angers de M. le Président de la République, le 1er décembre 2006.

Colloques

Braibant Jean-Marc, « Comment aider et accompagner les enseignants universitaires à améliorer la qualité des examens par QCM sur base des analyses d'items ? », Actes du 27e Colloque de l'ADMEE Europe, 2015, Liège.

Castel Robert, « L'expert mandaté et l'expert instituant », Actes du colloque du centre de recherche et d'étude sociologiques appliquées de la Loire, 1985.

Colloque du 28 août 2018, *Les grands défis de l'École dans une société mondialisée : de quelle École avons-nous besoin ?* Entretiens Enseignants-Entreprises 2018 - Les entreprises dans la mondialisation.

Ischinger Barbara, « What Do International Organizations have to do with Education ? », Actes du colloque *l'École apprend elle le monde ?*, 4 juillet 2014.

Mons Nathalie, *Policy Responses to PISA : a comparative investigation across OECD countries*, Keynote speech for the 14th International Conference of the Association for Educational Assessment, International Surveys, Policy Borrowing and National Assessment, Université de la Sorbonne, Paris, 2013.

Rocher Thierry, *Intervention au colloque du CSEN*, « Le rôle de l'expérimentation dans le domaine éducatif » 1er février 2018, Collège de France.

Roqueplo Philippe, *Entre savoir et décision, l'expertise scientifique*, conférence organisée par le groupe Sciences en questions, Paris, INRA, 9 avril 1996.

Sites internet

Site de l'INSEE sur le Service Statistique Ministériel : www.insee.fr/fr/information/2113161

Site du ministère de l'Éducation nationale sur les résultats CEDRE : www.education.gouv.fr/cedre-2013-grande-stabilite-des-acquis-en-sciences-en-fin-d-ecole-depuis-2007-8408

Site du ministère de l'Education nationale sur la Dgesco :
www.education.gouv.fr/cid978/la-direction-generale-de-l-enseignement-scolaire.html

Site du ministère de l'Education nationale sur la LOLF :
www.education.gouv.fr/pid23-cid31/la-lolf-est-que-est.html

Site de l'OCDE sur l'équipe PISA :
www.oecd.org/pisa/contacts/whoswhoinpisa.htm

Site de l'OCDE sur PIAAC : www.oecd.org/skills/piaac/

Site du Who's who France : www.whoswho.fr

Annexes

Annexe 1. Liste des entretiens réalisés

N° entretien	Nom de l'interlocuteur	Organisme	Fonction	Date	Durée	Lieu
1	Sophie Vayssettes	OCDE	responsable de projet PISA	5 novembre 2015	1h30	un café
2	Bruno Trosseille	DEPP	directeur du bureau des évaluations des élèves	5 janvier 2016	2h	son bureau
3	Bruno Trosseille	DEPP	directeur du bureau des évaluations des élèves	14 avril 2016	1h45	son bureau
4	Anne Laure Monnier	Conseil Supérieur des Programmes	Chargée de mission, ancienne agente de la DEPP	12 septembre 2016	2h	un café
5	Agathe Cagé	Ministère de l'Education nationale	conseillère de Najat Vallaud Belkacem (et avant de Vincent Peillon) en charge du second degré, des programmes et de l'évaluation des élèves	octobre 2016	1h15	son bureau
6	Marie Christine Blandin	Sénat	Sénatrice du Nord, ancienne Présidente de la Commission Education et Culture du Sénat	3 novembre 2016	1h	son bureau
7	Eric Charbonnier	OCDE	Responsable communication Direction de l'éducation de l'OCDE	16 décembre 2016	1h	l'OCDE

8	Collaborateur parlementaire*	Sénat	Groupe de gauche - Commission éducation et affaires culturelles	18 décembre 2016	45 minutes	Sénat
9	Collaboratrice parlementaire*	Assemblée nationale	Groupe de gauche - Commission éducation et affaires culturelles	12 mars 2017	1h	un café
10	Andreas Schleicher	OCDE	Directeur de la Direction de l'éducation de l'OCDE	16 janvier 2017	1h	son bureau
11	Agent OCDE*	OCDE	Analyste Direction de l'éducation de l'OCDE	2 février 2017	2h	un café
12	Isabelle Robin	DGESCO	Directrice du département de la recherche et du développement, de l'innovation et de l'expérimentation :	7 avril 2017	1h	son bureau
13	Françoise Cartron	Sénat	Sénatrice PS de Gironde, rapporteure de la loi refondation de l'école,	10 mai 2017	1h	son bureau
14	Collaborateur parlementaire*	Sénat	Groupe de gauche - Commission éducation et affaires culturelles	12 mai 2017	45 minutes	un café

15	Yves Durand	Assemblée nationale	Sénateur PS puis LREM de Mayenne, rapporteur de la loi refondation de l'école, président du comité de suivi et membre du comité d'évaluation des politiques publiques	14 juin 2017	2h	son bureau
16	Agente OCDE*	OCDE	Statisticienne Direction de l'Education	19 juin 2017	1h	un café
17	Franck Salles	DEPP	NPM et chargé plus spécialement des rapports en science et maths	11 juillet 2017	1h30	son bureau
18	Jean Pierre Jeantheau	DEPP	NPM	18 mars 2019	1h30	skype
19	Jean Pierre Jeantheau	DEPP	NPM	21 mars 2019	1h	skype
20	Catherine Moisan	DEPP	Directrice de la DEPP 2013-2015	18 juillet 2019	1h	téléphone
21	Ronan Vourc'h	DEPP	Adjoint chef du bureau de l'évaluation	10 juin 2020	1h	téléphone
22	Collaborateur parlementaire*	Sénat	Groupe de gauche - Commission éducation et affaires culturelles	22 juin 2020	1h	skype
23	Agente de la DEPP	DEPP	Statisticienne	1er juillet 2020	1h	téléphone

Annexe 2. Note anonyme rédigée en interne à la DEPP à propos de PISA

Critiques et enjeux des enquêtes internationales sur les acquis des élèves : à propos de l'enquête PISA.

En réponse à votre demande, je présente dans cette note brève et schématique quelques réflexions qui, demanderaient à être davantage argumenté, A) sur les critiques méthodologiques, ensuite B) sur les enjeux politiques et économiques des enquêtes internationales.

A. Critique des enquêtes internationales:

La critique des enquêtes internationales est délicate d'une part, car leur utilité n'est pas contestable - il faut les faire -, et d'autre part, car leurs méthodologies sont complexes de ce fait, la critique ne peut se faire qu'à un niveau souvent très technique. Les méthodologies sur lesquelles reposent ces enquêtes sont présentées comme ayant été l'objet de travaux scientifiques rigoureux et peu contestables. La quantité et la qualité de ces travaux est certaine, néanmoins tous les problèmes sont encore loin d'être résolus. Par ailleurs, pour des raisons institutionnelles, les méthodes adoptées sont, principalement, celles qui ont été développées aux Etats Unis pour des comparaisons fédérales. Les organismes travaillant à ces évaluations fédérales ont cherché à étendre leur champ d'action aux comparaisons internationales. On peut se demander dans quelle mesure ce cadre peut être valablement étendu aux comparaisons internationales. Par voie de conséquence, les travaux d'autres pays, d'autres organismes, d'autres chercheurs ont été soit laissés de côté soit minimisés. Il faut contrairement au discours que tiennent ces organisations se garder de considérer que la méthode adoptée est la seule et la meilleure. On peut ici s'interroger sur deux points.

I. Leurs aspects réducteurs :

1. Les enquêtes internationales sont obligées de tenir compte de nombreuses contraintes sur le plan méthodologique. On peut dire qu'on mesure là où on sait mesurer (mettre les pays sur une seule échelle) plutôt que de chercher à évaluer ce qui nous intéresse (faire ressortir les particularités nationales des apprentissages).
2. Les épreuves devant être adaptées à de nombreux pays sont, de ce fait, décontextualisées, c'est à dire indépendantes des programmes nationaux. Les performances des élèves aux exercices peuvent donc être très variables selon leur degré de familiarité avec les exercices.
3. La présentation des résultats privilégie largement l'unidimensionnalité (tous les pays placés sur une échelle unique) alors que l'univers des compétences serait plutôt multidimensionnel.

II. Le respect des cultures:

1. Bien que chaque pays puisse proposer des exercices, la majorité des épreuves est cependant d'origine anglo-saxonne (voire reprises des épreuves d'évaluation des jeunes américains) ou sinon calquées sur des modèles anglo-saxons.

2. Les méthodologies pour adapter (traduire) tests et questionnaires se sont largement développées depuis une vingtaine d'années. néanmoins, ces méthodologies ne sont pas entièrement pertinentes dès lors que les épreuves portent essentiellement sur un matériel verbal complexe (cas de la littérature). On fait comme si la difficulté de compréhension d'un texte était la même dans toutes les langues.
3. Il faut également signaler que les méthodes d'identification des biais culturels ne permettent pas de les identifier tous lorsqu'on travaille sur des épreuves traduites car ces méthodes sont fiables pour travailler sur des épreuves dans une même langue.

B. Des enjeux politiques et économiques:

Les enquêtes internationales sur le rendement des élèves se sont mises en place au cours des années soixante sous l'égide de l'IEA. La France a largement participé dès le départ à ces travaux. La prise en main par l'OCDE d'une partie de ces enquêtes et l'arrivée d'institutions nouvelles a mis en évidence les enjeux politiques et économiques de la maîtrise d'œuvre de ces enquêtes. Il s'agit du marché de l'évaluation et en arrière plan de celui de l'éducation puisque l'évaluation permet de juger de la qualité des formations dispensées voire de les certifier. A titre d'exemple citons les faits suivants :

- Un organisme comme ETS (*Educational Testing Service*), organisme privé américain, qui s'est construit un empire politique et commercial sur les tests d'admission dans les universités américaines et auteur du TOEFL (*Test Of English as Foreign Language*) cherche maintenant à développer une politique commerciale internationale (ETS a ouvert une représentation en Europe).
- On assiste à un phénomène de monopole et de renvoi d'ascenseur entre institutions et personnes. Les institutions sont présentes dans les différents organismes : ainsi la Commission Internationale des Tests (ONG) a élaboré des recommandations sur l'adaptation et la traduction des tests. Le colloque au cours duquel ont été présentées ces recommandations était animé par la Commission Internationale des Tests mais aussi par ETS avec la participation de Microsoft. A ce colloque intervenaient des universitaires à la fois membre de la commission des tests et participants à PISA. Les recommandations garantissent que les produits satisfont à des normes de qualité, on voit donc le danger de la collusion entre une organisation professionnelle garante de la qualité des instruments d'évaluation et un éditeur de test.
- Il est important de dire que ces enquêtes internationales, sont une vitrine pour un certain nombre d'organisations afin d'apparaître comme les meilleures, voire les seules, dans ce domaine. Les mêmes institutions ont fait plusieurs tentatives pour se placer auprès de l'UNESCO. On peut, il me semble, faire l'hypothèse - sans que cela apparaisse pour de la paranoïa antimondialisation et anti-américaine - qu'il est en train se développer un impérialisme culturel dans ce domaine de l'évaluation. Signalons qu'il existe déjà largement dans le domaine de l'évaluation psychologique puisque - pour parler rapidement - on mesure "l'intelligence" des petits français avec des adaptations de tests américains et la personnalité des cadres français avec des questionnaires de personnalité adaptés de questionnaires américains.

Quelle réponse apporter à ce mouvement ?

Nous avons pensé qu'il était nécessaire d'affirmer une présence française dans ce domaine. Ceci se traduit d'abord par une réponse à l'appel d'offre de l'OCDE pour l'enquête PISA (candidature non retenue mais qui a permis de montrer une présence européenne coordonnée par la DPD). Par ailleurs, en s'appuyant sur le réseau européen des responsables

des politiques d'évaluation des systèmes éducatifs (coordonné par Gérard Bonnet de la DPD), la DPD a piloté plusieurs programmes d'enquêtes européennes sur les comparaisons du niveau des élèves en anglais et en lecture entre pays européens volontaires (programmes Socrates).

Pour terminer, je dirai que les systèmes d'évaluation ne sont pas indépendants des systèmes de valeurs et des contextes culturels. La France a une longue tradition dans ce domaine et son système d'évaluation, malgré des défauts soulignés par les récents rapports au HCEE, est reconnu au plan international comme une "expérience exemplaire" (phrase de G. Landsheere, universitaire belge, ancien Président de l'IEA). Notre approche de l'évaluation est originale en particulier par l'accent mis sur l'évaluation diagnostique et l'implication de tous les acteurs et partenaires du système éducatif au dispositif d'évaluation. Il importe dans ce domaine aussi de sauvegarder l'exception culturelle.

RÉSUMÉ

En 2000 est lancée l'évaluation PISA sous l'égide de l'OCDE, à laquelle participe la France. La DEPP, direction des évaluations statistiques interne au ministère de l'Éducation nationale, à l'institutionnalisation récente, est chargée d'en suivre la mise en œuvre pour la France et de garantir la validité des résultats pour aider au pilotage de la politique éducative nationale. Dans un contexte où l'évaluation des résultats de l'action éducative est encore peu instituée, l'irruption de PISA dans le panel des évaluations à disposition des décideurs publics remet en question le rôle de producteur de données pour l'action que la DEPP s'est tout juste vue confier. A la fois interlocutrice privilégiée de l'OCDE pour PISA mais sans réel pouvoir d'agir en profondeur sur les choix de méthode et d'objectifs faits par l'organisation internationale, en charge d'évaluations nationales encore peu utilisées pour piloter l'action publique, la DEPP va devoir repenser les caractéristiques de son expertise pour s'assurer que ses évaluations et ses analyses sont bien employées à cette fin, complétées par les données internationales de PISA et non l'inverse. Afin d'observer le repositionnement des experts de l'évaluation statistique en éducation au moment de la construction de PISA et jusqu'à la publication relative à son cycle 2016, cette thèse s'intéresse à la construction de l'identité d'expert en évaluation des résultats de l'action éducative de la DEPP et de l'OCDE avec PISA. Elle cherche ensuite à mettre en lumière les stratégies de légitimation de ses travaux d'évaluation menées par les agents de la DEPP, entre critique méthodologique de PISA et intérêt philosophique et technique de l'objet. Enfin, elle observe la stratégie à l'épreuve de l'usage, en regardant la façon dont les résultats PISA et les évaluations de la DEPP sont employées dans le cadre du contrôle de l'action du gouvernement et lors des deux grandes réformes de la période, l'instauration du socle commun de compétences en 2005 et la refondation de l'école pour lutter contre les inégalités scolaires en 2013.

MOTS CLÉS

Evaluation des politiques éducatives, organisation internationale, service statistique ministériel, expertise, PISA, OCDE, ministère de l'Éducation nationale, concurrence des acteurs, résultats de l'éducation, fabrique de la loi, pilotage de l'action publique, évaluation standardisée, entrepreneur de problème public.

ABSTRACT

In 2000, PISA is launched under the auspices of the OECD, with the participation of France. The DEPP, the internal statistical evaluation directorate of the Ministry of education, which had recently been institutionalized, is put in charge of monitoring its implementation in France and ensuring the validity of the results to help steer the national education policy. The irruption of PISA as a tool for policy makers to evaluate education policies in France, where the evaluation of education policies is in the process of being institutionalised, calls into question the recently acquired role of the DEEP as data producer. The Directorate is the OECD's direct interlocutor, but has no real ability to influence the objectives and choices of method made by the international organisation. The Directorate is at the same time in charge of the national education evaluations and assessments which are still little used to steer public action. The DEPP will thus have to rethink its expertise to ensure that its evaluations and assessments, as well as its analyses of the results, are used for policy definition, and are complemented by PISA data, not the other way around. In order to study the repositioning of the DEEP and the OECD as experts in the evaluation of education policy, this thesis analyses the process of construction of such role using PISA, from the construction of international assessment until its 2016 cycle. It then analyses the strategies for legitimising DEEP's role as education policy evaluator, highlighting the paradox of their methodological criticism of PISA and their philosophical and technical interest for it. Finally, it looks at that strategy in action, observing how PISA results and DEPP evaluations and assessments are used for the control of governmental action and in the two major reforms of that period, the establishment of the common base of skills in 2005 and the 2013 "refounding of the school" aiming at tackling school inequities.

KEYWORDS

Education policies assessment, international organisation, ministerial service of statistics, expertise, PISA, OECD, ministry of education, competition between actors, education outcomes, policy making, public action steering, standardised evaluation, public problem entrepreneur.