



HAL
open science

Évaluer autrement avec le numérique : coévaluation instrumentée pour la production d'écrit en français langue première au cycle 4 du collège français

Nicolas De Khovrine

► To cite this version:

Nicolas De Khovrine. Évaluer autrement avec le numérique : coévaluation instrumentée pour la production d'écrit en français langue première au cycle 4 du collège français. Education. Université Paul Valéry - Montpellier III, 2023. Français. NNT : 2023MON30058 . tel-04860619

HAL Id: tel-04860619

<https://theses.hal.science/tel-04860619v1>

Submitted on 1 Jan 2025

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

Pour obtenir le grade de
Docteur

Délivré par l'**université Paul-Valéry – Montpellier 3**

Préparée au sein de l'**École Doctorale 58**
Langues, littératures, cultures, civilisations
et des unités de recherche
LHUMAIN et LIRDEF

Spécialité : **Sciences de l'éducation et de la formation**

Présentée par **Nicolas de KHOVRINE**

**Évaluer autrement avec le numérique :
coévaluation instrumentée pour la production d'écrit
en français langue première au cycle 4 du collège
français**

Soutenue le 16 11 2023 devant le jury composé de :

M. Laurent LESCOUARCH, PR, Université de Caen	Rapporteur
M. Pascal MARQUET, PR, Université de Strasbourg	Rapporteur
Mme Marie-Paule JACQUES, PR, Université de Toulouse	Présidente du jury
Mme Stéphanie MAILLES VIARD METZ, PR, Aix-Marseille Université	Directrice
M. Frédéric TORTERAT, PR, Université de Montpellier	Codirecteur



« En regardant la terre de loin, vous réalisez qu'elle est trop petite pour qu'il y ait un conflit et juste assez grande pour une coopération. »

Youri Gagarine (1934-1968), cosmonaute, premier homme dans l'espace.

RESUME

Face à l'hétérogénéité des classes au collège, l'École française incite ses enseignants à pratiquer la différenciation pédagogique, mais cette approche se limite souvent à des dispositifs d'individualisation des apprentissages, susceptibles de conforter, malgré les efforts, une baisse du sentiment d'efficacité chez les élèves, notamment lors des évaluations.

Or selon le paradigme du collectif apprenant, les pédagogies coopératives permettent d'envisager une modalité selon laquelle apprenants et enseignants participent conjointement au processus évaluatif, à travers ce qu'il convient d'appeler la coévaluation. Une telle modalité évaluative est toutefois difficilement envisageable sans recourir à des outils numériques.

La présente recherche collaborative par méthodes mixtes vise à mesurer les effets d'un dispositif pédagogique numérique de coévaluation sur les apprentissages de collégiens en maîtrise de la langue et production d'écrit en français langue première et sur leur sentiment d'efficacité personnelle.

Conçu selon un détournement d'usage du logiciel libre Moodle Atelier, ce dispositif a été expérimenté à 9 reprises en 2021-2022 avec 80 élèves issus de 4 classes de cycle 4 (5^e, 4^e, 3^e), dans 3 collèges différents, dans les académies de Paris et Montpellier.

Un algorithme rend compte de la performance de chaque élève, comme scripteur et comme évaluateur, selon l'approche par compétences. Le sentiment d'efficacité personnelle des élèves participants est mesuré avant et après chaque itération du protocole semi expérimental, tandis que des données qualitatives issues de l'observation directe et indirecte sont recueillies pour éclairer la genèse instrumentale de ce dispositif de coévaluation médiatisée.

Les résultats indiquent une augmentation significative du sentiment d'efficacité personnelle des élèves participants, ainsi que de leurs performances comme scripteurs et comme évaluateurs. Un nombre substantiel de conflits instrumentaux se sont toutefois manifestés, accompagnés de phénomènes de régulation dans la plupart des cas : une typologie est proposée pour analyser ces dynamiques de conflictualité/régulation et éclairer les usages.

MOTS-CLES

pédagogies coopératives ; didactique du français ; coévaluation ; numérique éducatif ; approche instrumentale ; textométrie ; méthodes mixtes ; recherche collaborative

ABSTRACT

Faced with the heterogeneity of middle school classes, the French school system encourages its teachers to practice pedagogical differentiation, but this approach is often limited to individualizing learning, which, despite efforts, can lead to a decline in students' sense of effectiveness, particularly when it comes to assessment. However, according to the paradigm of the collective learner, cooperative pedagogies make it possible to envisage a modality in which learners and teachers participate jointly in the evaluation process, through what is known as co-assessment. However, such an evaluative modality is difficult to implement without recourse to digital tools. The aim of this mixed-methods collaborative research project is to measure the effects of a digital co-assessment system on the learning and self-efficacy of middle-school students in language proficiency and writing production in French as a first language.

Conceived as a hijack of the open-source Moodle Atelier software, this system was tested 9 times in 2021-2022 with 80 pupils from four cycle 4 classes (5e, 4e, 3e), in three different Middle Schools, in the Paris and Montpellier academies.

An algorithm reports on each student's performance, both as a writer and as an evaluator, according to the competency-based approach. Participating students' sense of self-efficacy is measured before and after each iteration of the semi-experimental protocol, while qualitative data from direct and indirect observation are collected to shed light on the instrumental genesis of this mediated co-evaluation device.

The results indicate a significant increase in participating students' sense of self-efficacy, as well as in their performance as scribblers and evaluators. However, a substantial number of instrumental conflicts arose, accompanied by regulatory phenomena in most cases: a typology is proposed to analyze these conflict/regulation dynamics and shed light on usage.

KEYWORDS

Collaborative learning; Didactics of French; Co-assessment; Computer-supported collaborative learning (CSCL); Instrumental approach; Textometric analysis; Mixed methods; Community-based participatory research

REMERCIEMENTS

Mes plus vifs remerciements vont à ma directrice de thèse, Stéphanie Mailles Viard Metz et à mon codirecteur Frédéric Torterat. C'est grâce à votre confiance et à vos encouragements que j'ai pu mener ce travail à bien. Votre disponibilité, votre qualité d'écoute et l'acuité de vos conseils ont été des aides précieuses tout au long de ce projet : soyez-en ici chaleureusement remerciés.

Je tiens également à exprimer ma profonde gratitude aux membres du jury : Madame Marie-Paule Jacques, Monsieur Laurent Lescouarch et Monsieur Pascal Marquet, qui honorent mon travail de leur lecture et de leur présence.

Surtout, cette étude n'aurait pu aboutir sans Claire Lavit-Bérard, Gaëlle-Andrée Morois et Pascale Brulin, les trois excellentes professeures de Lettres modernes avec qui j'ai eu l'honneur de vivre cette recherche collaborative. Vous avez considérablement enrichi cette étude grâce à votre engagement, votre expertise professionnelle et votre sens du travail en équipe. Je tiens tout particulièrement à remercier vos élèves de 5e, 4e et 3e pour leur implication enthousiaste dans ce projet : par l'entremise de la communauté éducative et scientifique, j'espère qu'il pourra contribuer aux apprentissages d'autres élèves après eux. Je voudrais en outre exprimer ma reconnaissance à Alix Frediani, professeure-documentaliste pour sa généreuse et précieuse contribution.

Un grand merci à mes précieux interlocuteurs des Cellules Académiques Recherche Développement Innovation Expérimentation (CARDIE) et des délégations académiques du numérique pour l'éducation (DANE) de Montpellier et Paris pour l'accueil qu'ils ont réservé au projet, et pour l'avoir rendu administrativement et techniquement possible : merci à Laurence Grandet (CARDIE 34), Farida Pariolleau (CARDIE 75), Jimmy Baret (DRANE Occitanie) et Delphine Stucchi (DANE 75). Je tiens également à exprimer ma sincère gratitude aux IA-IPR de Lettres de Montpellier et Paris, qui ont validé le projet, et tout particulièrement à MM. Frédéric Miquel et Jean-Philippe Taboulot pour leur soutien précieux. Ma profonde reconnaissance va en outre aux chefs d'établissement qui ont bien voulu donner sa chance à ce projet expérimental et qui en ont facilité la mise en œuvre : merci à Fabienne Dondard, Issa Sène et Guillaume Gouchault.

A Pierre, José, Nihal et Amadou, mes collègues doctorants au sein de l'équipe *Human Connect* : grâce à la belle dynamique d'entraide insufflée par notre directrice commune Stéphanie Mailles Viard Metz, vous avez été mes précieux compagnons de route pendant ces

cinq années : je tiens à vous témoigner ici toute mon amitié et ma reconnaissance. Mes remerciements vont en outre aux équipes des laboratoires LHUMAIN et LIRDEF, ainsi qu'à l'école doctorale 58, et tout spécialement Anne Le Perrot à la formation doctorale. Un grand merci par ailleurs aux bibliothécaires de la CRD et de la BIU à Montpellier, ainsi qu'à Yann Hellebaque, à Jérôme Marini et Florian Boulland de l'équipe d'appui à la science ouverte de l'Université de Montpellier pour leur grande aide.

Merci aux *Cahiers Pédagogiques* pour la publication de notre « article à huit mains ».

Toute ma gratitude à mes relecteurs : mes collègues Claire, Gaëlle et Pascale, et surtout un immense merci à Pierre.

Je tiens enfin à exprimer ma profonde et affectueuse gratitude à l'égard de tous ceux qui m'ont infailliblement soutenu et accompagné avec sollicitude durant ces cinq années, au premier rang desquels mon épouse Marie, ma fille Olga, ma famille, ma belle-famille et mes amis.

Table des matières

<i>RESUME</i>	3
<i>MOTS-CLES</i>	3
<i>ABSTRACT</i>	4
<i>KEYWORDS</i>	4
<i>REMERCIEMENTS</i>	5
<i>INTRODUCTION</i>	17
<i>PARTIE I : ETAT DE L'ART</i>	22
<i>1 Le défi des hétérogénéités au collège</i>	22
1.1 Origines du « collège unique » en France.....	22
1.2 Massification scolaire vs. démocratisation de la réussite	22
1.3 Hétérogénéités contextuelles et structurelles du collège en France.....	24
1.4 L'hétérogénéité en milieu scolaire : définition et contextualisation.....	25
Ce qu'il faut retenir	26
<i>2 Vers une opérationnalisation de la différenciation pédagogique</i>	26
2.1 Une approche privilégiée par l'institution.....	26
2.2 Une approche difficile à opérationnaliser	27
2.3 La différenciation pédagogique en pratique	29
2.4 Limites de la différenciation pédagogique.....	32
Ce qu'il faut retenir	34
<i>3 Les pédagogies coopératives</i>	34
3.1 Coopération vs. collaboration : tentative de désambiguïsation.....	34
3.2 Origines et place dans les textes officiels	37
3.3 Approches contemporaines.....	38
Ce qu'il faut retenir	42
<i>4 L'autodirection des apprentissages</i>	42

4.1	L'autorégulation des apprentissages	42
4.2	L'auto-efficacité, « coefficient » de l'autodirection des apprentissages.....	50
4.3	Autodirection et apprentissages de préadolescents	55
	Ce qu'il faut retenir	57
5	<i>Vers une évaluation coopérative.....</i>	<i>57</i>
5.1	Les quatre finalités évaluatives	57
5.2	Les trois principales modalités évaluatives	59
5.3	La coévaluation.....	62
	Ce qu'il faut retenir	64
6	<i>La coévaluation numérique individualisée</i>	<i>65</i>
6.1	La genèse instrumentale.....	65
6.2	Le conflit instrumental en EIAH.....	66
6.3	Artéfacts numériques pour la coévaluation	69
	Ce qu'il faut retenir	77
7	<i>Didactisation de la production écrite en EIAH.....</i>	<i>78</i>
7.1	Production écrite et étude de la langue au collège.....	78
7.2	La production écrite sur traitement de texte	78
7.3	Production écrite sur traitement de texte et inter-évaluation.....	79
7.4	Dispositifs numériques pour l'inter-évaluation de productions écrites.....	79
	Ce qu'il faut retenir	80
8	<i>La recherche par méthodes mixtes (RMM)</i>	<i>80</i>
8.1	Une approche adaptée pour l'étude de la différenciation pédagogique	80
8.2	Un concept polysémique ?.....	81
8.3	Stratégies pour l'intégration des données	82
8.4	Modèles (ou devis) principaux.....	82
	Ce qu'il faut retenir	83
9	<i>La recherche collaborative</i>	<i>83</i>
9.1	Une autre approche adaptée à l'étude de la DP ?.....	83

9.2	Une posture épistémologique singulière	84
9.3	Un métissage des savoirs	85
9.4	Une dimension communautaire.....	85
	Ce qu'il faut retenir	86
<i>PARTIE II : PROBLEMATIQUE ET HYPOTHESES</i>		87
1	<i>Problématique</i>	87
2	<i>Variables générales</i>	88
3	<i>Hypothèse générale</i>	89
4	<i>Variables opérationnelles</i>	89
4.1	Indicateurs en lien avec la théorie instrumentale	89
4.2	Indicateur en lien avec la production écrite sur traitement de texte	90
4.3	Indicateurs en lien avec l'autodirection des apprentissages	90
4.4	Observable en lien avec le paradigme de la recherche collaborative.....	91
5	<i>Situation choisie</i>	91
6	<i>Prédictions</i>	93
6.1	HO 1 : apprentissages en français langue première	93
6.2	HO 2 : autodirection des apprentissages	93
6.3	HO 3 : apprentissages coopératifs.....	94
6.4	HO 4 : conflits et régulations instrumentaux en EIAH.....	94
6.5	Prédictions relatives à la fiabilité des instruments	94
<i>PARTIE III : CHOIX METHODOLOGIQUES</i>		96
1	<i>Approche épistémologique</i>	96
1.1	Une expérimentation semi-contrôlée en situation écologique.....	96
1.2	Une recherche-intervention collaborative	97
1.3	Une recherche par méthodes mixtes (RMM)	97
2	<i>Participants</i>	98
2.1	Collèges.....	99

2.2	Académies et partenaires institutionnels afférents	99
2.3	Enseignantes partenaires	99
2.4	Échantillons.....	100
3	<i>Protocole de recherche.....</i>	101
3.1	Principales caractéristiques	101
3.2	Artéfact didactique : production écrite	103
3.3	Artéfact pédagogique : coévaluation.....	114
3.4	Artéfact numérique : Moodle Atelier	115
3.5	Durée et nombre d'itérations.....	118
4	<i>Recueil et traitement des données</i>	118
4.1	Questionnaire en ligne.....	119
4.2	Traitement des données numériques	131
4.3	Focus groupe (échantillon C')	145
4.4	Retours écrits sur expérience (échantillon C)	145
4.5	Focus groupe (échantillon C).....	147
4.6	Entretiens semidirigés et focus groupe (enseignantes partenaires)	147
5	<i>Apparatus</i>	149
5.1	Moodle	149
5.2	Ordinateurs en salle informatique	150
5.3	Traitement de texte utilisé pour les productions écrites	150
5.4	Logiciel utilisé pour le questionnaire en ligne.....	150
5.5	Logiciel utilisé pour les analyses statistiques	152
5.6	Logiciel utilisé pour les analyses textométriques	152
5.7	Ordinateur utilisé pour la présente recherche.....	154
5.8	Entrepôt de données	154
6	<i>Aspects déontologiques</i>	155
	<i>PARTIE IV : RESULTATS.....</i>	156
1	<i>Compétences en français (HO 1).....</i>	157

	1.1	Augmentation modérée du score de scripteur de SE sur T2-T4	157
T6	1.2	Différence significative entre les scores de scripteurs de l'échantillon C sur T2-158	
	1.3	Augmentation du SEP(scol) de SE sur T3-T4.....	159
	1.4	Discussion (HO 1)	160
	2	<i>Autodirection des apprentissages (HO 2)</i>	161
	2.1	Autorégulation des apprentissages (HO 2a).....	161
	2.2	Métacognition (HO 2b)	164
	2.3	Sentiment d'efficacité personnelle (HO 2c).....	167
	2.4	Discussion (HO 2)	171
	3	<i>Apprentissages coopératifs (HO 3)</i>	173
	3.1	SEP(coop)	174
	3.2	SEC-IAP	175
	3.3	Corrélations entre SEP(coop) et SEC-IAP.....	176
	3.4	Discussion (HO 3)	177
	4	<i>Conflits et régulations instrumentaux (HO 4)</i>	180
	4.1	Situations d'apprentissage (HO 4a)	181
	4.2	Situations d'enseignement (HO 4b)	200
	4.3	Discussion (HO 4)	225
		<i>PARTIE V : DISCUSSION GENERALE</i>	226
	1	<i>Coévaluation numérique, maîtrise de la langue et production d'écrit</i>	226
	1.1	Contresens pédagogique et empêchement didactique	227
	1.2	Prérequis didactiques et conflit instrumental pédagogique.....	228
	1.3	Problèmes de discernement docimologique	231
	1.4	Combinatoire conflictuelle en lien avec les interfaces de l'artéfact numérique	234
didactique	1.5	Synergie régulatrice favorisant la motivation intrinsèque, en lien avec l'artéfact	235

1.6	Le dispositif pédagonumérique : source à la fois de conflits et de régulations didactiques	237
1.7	Lecture et coévaluation individualisée outillée par le numérique.....	239
1.8	Ajustements didactiques d'une itération à l'autre et effets sur la maîtrise des compétences cibles	240
1.9	Artéfact numérique et prolongements possibles pour l'artéfact didactique .	242
	Ce qu'il faut retenir	243
2	<i>Coévaluation numérique et autodirection des apprentissages</i>	244
2.1	Coévaluation numérique et autorégulation des apprentissages	245
2.2	Coévaluation numérique, métacognition et sentiment d'efficacité personnelle	247
	Ce qu'il faut retenir	252
3	<i>La coévaluation numérique est-elle coopérative ?</i>	252
3.1	Coopération entre élèves et enseignants	253
3.2	Coopération entre élèves	254
3.3	Relation d'aide	255
3.4	Souci de formuler un écrit commentatif adapté	255
3.5	Travail de groupe et entraide	257
	Ce qu'il faut retenir	258
4	<i>Coévaluation numérique et « dialectique instrumentale »</i>	259
4.1	Combinatoires conflictuelles et synergies régulatrices.....	260
4.2	Des fonctionnalités régulatrices, apanage de l'artéfact numérique	263
4.3	Peut-on pratiquer la coévaluation individualisée sans artéfact numérique ?	264
4.4	Degrés de conflictualité et effets sur les apprentissages	266
4.5	Retour réflexif sur le point de vue des enseignantes partenaires	268
	Ce qu'il faut retenir	271
5	<i>Coévaluation numérique et différenciation pédagogique</i>	272
5.1	Artéfact numérique et différenciation pédagogique.....	272

5.2	Différenciation pédagogique et métacognition.....	275
5.3	Coévaluation numérique et inégalités socioscolaires : un bilan nuancé mais encourageant	276
	Ce qu'il faut retenir.....	286
6	<i>La recherche collaborative comme retravail de l'empirie.....</i>	287
	Ce qu'il faut retenir.....	290
7	<i>Limites de cette étude.....</i>	292
	Ce qu'il faut retenir.....	298
	<i>CONCLUSION.....</i>	299
1	<i>Bilan : réponse à la question de recherche.....</i>	299
2	<i>Perspectives pour la recherche et l'enseignement.....</i>	303
	<i>REFERENCES BIBLIOGRAPHIQUES.....</i>	308
	<i>ANNEXES.....</i>	336
1	<i>Résultats détaillés.....</i>	336
1.1	Résultats préliminaires relatifs à la validité des données.....	336
1.2	Hypothèse HO 1.....	342
1.3	Hypothèse HO 2.....	344
1.4	Hypothèse HO 3.....	351
1.5	Hypothèse HO 4.....	360
1.6	Etayage et performances des élèves : analyse transversale.....	368
2	<i>Matériel pédagogique des enseignantes partenaires.....</i>	370
2.1	T1/T2.....	370
2.2	T3/T4 : classe C (5 ^e).....	372
2.3	T5/T6 : classe C (5 ^e).....	373
2.4	Année N+1.....	374
3	<i>Questionnaire en ligne.....</i>	375
3.1	Information RGPD.....	375

3.2	Items d'identification	376
3.3	Items sur les habitudes évaluatives	376
3.4	Items de calibrage.....	376
3.5	Items SEC.....	376
3.6	Items SEPE.....	376
3.1	Correspondances items/dimensions SEPE	378
3.1	Codage, dimensions et libellé des items SEPE	381
3.2	Sous-échelle SEP(scol).....	385
3.3	Sous-échelle SEP(srl).....	386
3.4	Sous-échelle SEP(coop)	387
3.5	Items CI-EIAH.....	387
3.6	Items acceptabilité du questionnaire	387
3.7	Item libre	387
3.8	Captures d'écran	388
4	<i>Verbatim des retours écrits sur expérience (échantillon C, T6, N = 23)</i>	<i>389</i>
5	<i>Transcripts des focus groupes (C, T6).....</i>	<i>399</i>
5.1	Transcript du 1 ^{er} focus groupe (C1, T6, N = 12).....	399
5.2	Transcript du 2 nd focus groupe (C2, T6, N = 11)	402
6	<i>Entretiens semidirigés enseignantes (liens pérennes renvoyant vers les enregistrements audios).....</i>	<i>405</i>
6.1	Enseignante A (T2).....	405
6.2	Enseignante B (T4)	405
6.3	Professeure-documentaliste du collège C (T4).....	405
6.4	Enseignante C (T6).....	405
6.5	Enseignante B (année N+1).....	406
7	<i>Focus groupe enseignantes A, B et C (T7) : lien pérenne renvoyant vers l'enregistrement audio</i>	<i>406</i>
8	<i>Checklist Moodle Atelier.....</i>	<i>407</i>

8.1	Préparation par l’enseignante et les élèves.....	407
8.2	Configuration de l’atelier en amont.....	407
8.3	Manipulations sur l’atelier en aval	407
9	<i>Lettre type d’information aux parents.....</i>	408
10	<i>Article coécrit avec les 3 enseignantes partenaires.....</i>	409
	<i>TABLE DES ILLUSTRATIONS</i>	412

*Ce manuscrit prend en compte les **rectifications de l'orthographe de 1990**, préconisées par l'Académie française d'après le rapport du Conseil supérieur de la langue française.*

INTRODUCTION

L'évaluation est omniprésente dans nos sociétés modernes. Au quotidien, en entreprise ou sur les réseaux sociaux, tout est passé au crible : collègues de travail, magasins, restaurants... A l'école en particulier, « l'évaluation peut être considérée comme un acte politique, aux implications à la fois sociales et pédagogiques » (Gallego *et al.*, 2017 ; Murillo & Hidalgo, 2015; cités par Quesada *et al.*, 2019, p. 1, traduction libre)¹. Les implications pédagogiques de l'évaluation scolaire renvoient à la manière dont une compétence est évaluée, ce qui détermine en retour l'approche utilisée pour la mise en œuvre des situations d'enseignement/apprentissage. Pour leur part, les implications sociales de l'évaluation à l'école ont trait à la certification des compétences instituées par le curriculum : il s'agit de garantir qu'à la fin de la scolarité obligatoire chaque élève aura acquis une maîtrise suffisante des compétences du Socle commun de connaissances, de compétences et de culture². L'objectif visé est de donner à chacun les ressources nécessaires à un épanouissement personnel et professionnel, notamment dans une logique d'apprentissage tout au long de la vie. A cet égard, (Morin, 1999) met l'accent sur la nécessité de favoriser l'autonomie de l'esprit et l'autodidactisme chez les apprenants. Singulièrement, dans un contexte où la pandémie de Covid-19 a considérablement accéléré la diffusion du numérique éducatif (Caron, 2020; Croze, 2021; Maity *et al.*, 2021; Taglietti *et al.*, 2021), ces concepts d'autonomie de l'esprit et d'autodidactisme pourraient constituer un enjeu majeur. En effet, avec la diffusion des moyens numériques à l'école – qu'ils soient matériels ou logiciels – et face à la concurrence grandissante des algorithmes et de l'intelligence artificielle, il semble que des compétences comme l'esprit critique, la créativité, la collaboration ou l'intelligence collective (Bellet, 2022) revêtent une importance toute particulière.

À la question de l'autonomie de l'apprenant, on voit ainsi s'adjoindre une deuxième notion, connexe et tout aussi fondamentale pour l'apprentissage tout au long de la vie : la

¹ Les références sont citées selon la norme APA 7^e édition. Le format (Auteur, Année) employé tout au long du manuscrit est lié à l'utilisation d'un logiciel tiers de gestion automatisée des références bibliographiques.

² Le socle commun de connaissances, de compétences et de culture présente ce que tout élève doit savoir et maîtriser à 16 ans. Il rassemble l'ensemble des connaissances, compétences, valeurs et attitudes nécessaires pour réussir sa scolarité, sa vie d'individu et de futur citoyen. [...] La maîtrise du socle est nécessaire pour obtenir le diplôme national du brevet (DNB). (Source : <https://www.education.gouv.fr/le-socle-commun-de-connaissances-de-competences-et-de-culture-12512>, adresse URL consultée le 10/09/2023).

coopération (ou collaboration³). En effet, grâce aux apports de (Piaget, 1975), on sait que c'est la réalisation d'activités qui permet d'apprendre : c'est la théorie constructiviste de l'apprentissage (Kerzil, 2009) ; cependant (Vygotsky, 1985) ajoute à cette perspective une dimension sociale. Son approche socioconstructiviste met l'accent sur l'importance de la coopération/collaboration pour les apprentissages : ces activités gagnent à être réalisées en interaction avec des pairs plus avancés ou des adultes, dans le cadre de ce que le psychologue russe appelle la « zone proximale de développement », et qu'il définit comme « la disparité entre l'âge mental, ou le niveau de développement présent, qui est déterminé à l'aide des problèmes résolus de manière autonome, et le niveau qu'atteint l'enfant quand il résout des problèmes non plus tout seul mais en collaboration » (Vygotsky, 1985, p. 276). D'emblée, on voit ainsi apparaître la notion d'apprentissage vicariant – *i.e.* par imitation –, et l'importance du groupe pour les apprentissages selon Vygotski (Grangeat, 2016). Dans le prolongement de cette approche socioconstructiviste, la théorie de l'agentivité humaine (Bandura, 2011) nous enseigne que les notions d'autonomie et de coopération sont étroitement liées à une troisième notion, qui conditionne en quelque sorte l'émergence des deux autres : la confiance (Samaniego Cho *et al.*, 2021). Cette dernière a notamment été opérationnalisée par (Bandura, 2007) à travers le concept de « sentiment d'efficacité personnelle » (ou auto-efficacité).

Or « les thèmes de la liberté, de la responsabilité et de l'autonomie pour apprendre traversent l'histoire des idées pédagogiques, tant pour ce qui est de l'éducation initiale que pour ce qui concerne ce que l'on appelle aujourd'hui la formation tout au long de la vie » et « on peut même s'essayer à considérer l'éducation "nouvelle" comme aux sources de l'idée d'autoformation. Les précurseurs (Rousseau, Kant, Condorcet) et les grands maîtres de la pédagogie (Jacotot, Pestalozzi, Cousinet, Freinet, etc.) ont toujours porté au fronton de leurs systèmes pédagogiques l'autonomie du sujet apprenant » (Poisson *et al.*, 2015, pp. 121-122). Ainsi, avec Carré (Poisson *et al.*, 2015), et en lien avec les pédagogies de la coopération issues du courant de l'éducation nouvelle, on peut envisager l'autonomie de l'apprenant comme le principe même d'une autodirection des apprentissages, *i.e.* de l'articulation qui peut être opérée entre autorégulation et auto-efficacité. Peu à peu, on peut chercher à élaborer ainsi une « synthèse conceptuelle et opératoire » (Poisson *et al.*, 2015, p. 121) avec la confiance, la coopération et l'autonomie comme fondements pour les apprentissages scolaires (Algan, Huillery, *et al.*, 2018).

³ Dans la littérature scientifique, ces deux termes sont utilisés de diverses manières selon les auteurs. À ce stade nous les utilisons indifféremment, cependant nous proposerons une désambiguïsation, *infra*.

Mais dans quelle mesure un tel construit théorique serait-il adapté aux contingences du terrain ? À fortiori, à un terrain comme celui du collège en France, dont on sait qu'il est fortement confronté à la problématique des inégalités socioscolaires (Cayouette-Remblière & Moulin, 2019; Duru-Bellat & Mingat, 1988, 1997) ? En effet, bien qu'en 2022 l'Éducation nationale demeure le premier poste budgétaire de l'État⁴, depuis une trentaine d'années un nombre significatif de rapports publics, d'enquêtes internationales et de publications scientifiques tendent à mettre en lumière un accroissement des inégalités dans le système scolaire français (Cnesco, 2016; Delahaye, 2015). Ainsi, selon l'enquête PISA⁵ de 2018 (OCDE, 2019b) :

« Comme déjà observé lors des éditions précédentes du PISA, la France est l'un des pays de l'OCDE où le lien entre le statut socioéconomique et la performance dans PISA est le plus fort » (p. 1).

Monseur & Baye (2016, p. 9) confirment cette tendance, en affirmant que « la France figure parmi les pays les plus inéquitables – c'est-à-dire ceux où les résultats scolaires dépendent le plus de l'origine sociale – et ce, depuis le début des enquêtes PISA ».

Certains auteurs comme Rutkowski & Engel (2010), Bieber & Martens (2011) ou Duru-Bellat (2019) ont souligné les limites de la méthodologie et des préconisations de PISA. Pour autant, selon la Direction de l'évaluation, de la prospective et de la performance (DEPP), la France est en effet « le pays de l'OCDE où la performance [des élèves] est la plus liée au milieu social » (DEPP, 2021). Yebbou (2015, p. 33) note même « un accroissement préoccupant de l'influence des inégalités sociales sur les inégalités scolaires ».

Or si comme nous le disions plus haut la pandémie de Covid-19 a considérablement accéléré la digitalisation de l'enseignement (Daniel, 2020 ; Ewing, 2021 ; Maity *et al.*, 2021), elle a en même temps contribué à accentuer ces phénomènes d'inégalités socioscolaires (Azorín, 2020 ; Buisson-Fenet & Marx, 2022). Le collège, lieu marqué par les hétérogénéités (Duru-Bellat & Mingat, 1988) est particulièrement touché par ces phénomènes (Cayouette-Remblière & Moulin, 2019 ; Duru-Bellat & Kieffer, 1999, 2000 ; Prost, 2001). Face à ce

⁴ Avec même une hausse de 2,6% par rapport au budget de l'année précédente. Source : Projet de loi de finance adopté par l'Assemblée nationale pour 2022. Consulté le 26 juillet 2022 à l'adresse URL : <https://www.senat.fr/rap/121-163-314/121-163-3140.html#toc4>.

⁵ PISA : Programme international pour le suivi des acquis des élèves (« *Programme for International Student Assessment* »).

constat, le Conseil d'évaluation de l'École (CEE, 2022) indique dans son dernier rapport que « la prise en compte de l'hétérogénéité des élèves dans les pratiques professionnelles demeure une difficulté majeure soulevée par les établissements [...]. La gestion réussie de la différenciation [pédagogique], entre autres par la formation, est un outil solide d'équité scolaire, qui passe par une identification précise des besoins des élèves » (p. 33). Ce même rapport ajoute qu'il en va de même de « l'évaluation des élèves [...] qui constitue en quelque sorte le pendant évaluatif de la différenciation pédagogique. » D'où l'avantage qu'il y a, selon le CEE, « à s'interroger sur l'articulation entre pratiques pédagogiques et évaluation des élèves [...] » (CEE, 2022, p. 34).

En relation étroite avec l'évaluation scolaire, on voit ici émerger le concept de différenciation pédagogique, dont on sait qu'il reste difficile à définir tant il recouvre une variété de pratiques (Forget, 2017 ; Lescouarch, 2021). Ainsi, on peut se demander d'emblée **quelle(s) approche(s) épistémologique(s) serai(en)t les mieux à même d'opérationnaliser la différenciation pédagogique ?**

En lien avec la différenciation pédagogique, la question de l'évaluation est également posée : **comment mettre en œuvre avec des collégiens une évaluation qui favoriserait l'autodirection des apprentissages et la coopération entre élèves, pour mieux prendre en compte leur hétérogénéité ? La digitalisation de l'enseignement, source potentielle d'accroissement des inégalités socioscolaires (Azorín, 2020 ; Buisson-Fenet & Marx, 2022), pourrait-elle accompagner la mise en œuvre de telles pratiques évaluatives ?**

Le Conseil d'évaluation de l'école indique justement que : « financé par les collectivités, l'ENT⁶ et ses fonctionnalités, pas toujours exploitées, sont valorisés dans les rapports d'évaluation en tant que ressources [...]. L'intégration pédagogique de ces outils permet le renouvellement des pratiques, le travail collaboratif, l'évaluation par compétence et la différenciation, grâce à la diversification pédagogique [et grâce] au suivi personnalisé des élèves [et] de leurs apprentissages [...] » (CEE, 2022, p. 39). Pour répondre à ce déficit d'exploitation, par les enseignants, de ressources numériques pourtant disponibles et qui pourraient contribuer à une meilleure prise en compte de l'hétérogénéité des élèves, « un accent particulier est mis [par l'institution scolaire, *i.e.* par le Ministère chargé de l'Éducation

⁶ ENT : Espace numérique de travail. Il s'agit d'une plateforme accessible en ligne ou via l'intranet d'un établissement, agrégeant divers contenus et fonctionnalités destinés à l'apprentissage. Pour les collèges français, dans chaque région académique, un ENT local est piloté par la délégation régionale au numérique éducatif (DRANE).

nationale et localement par les rectorats d'académie] sur les personnes ressources à mobiliser (réfèrent [numérique], DRANE, sollicités pour des interventions sur site) » (CEE, 2022, p. 39). Parmi les personnes ressources mobilisables, **les chercheurs ne peuvent-ils pas contribuer significativement à l'appropriation de ressources numériques, e.g. en proposant d'expérimenter des dispositifs pédagognumériques dans le cadre de recherches-actions de type collaboratif, en partenariat avec des enseignants, praticiens de terrain ?**

Pour tenter de répondre à ces questions, plusieurs champs disciplinaires seront investis ici, comme la psychopédagogie ou l'ergonomie en EIAH, afin de nourrir un questionnement transversal. De fait, la présente recherche s'inscrit dans le domaine des Sciences de l'éducation et de la formation à travers la didactique du français langue première et les pédagogies coopératives, tout en s'intéressant à celui des Sciences du langage, à travers la didactique de la langue et les apprentissages linguistiques.

Après avoir présenté notre état de l'art (partie I), nous formulerons notre problématique et nos hypothèses (partie II). Nous détaillerons ensuite nos choix méthodologiques (partie III), puis nous présenterons et discuterons nos résultats hypothèse par hypothèse (partie IV), avant de proposer une discussion générale (partie V) et de conclure.

PARTIE I : ETAT DE L'ART

1 Le défi des hétérogénéités au collège

1.1 Origines du « collège unique » en France

En France, la décision prise en 1959 de prolonger l'obligation scolaire jusqu'à 16 ans correspond à une volonté politique de « moderniser le pays » et de lui « rendre son rang », en promouvant une « économie forte et créative, qui suppose une main-d'œuvre globalement mieux formée et des techniciens et chercheurs beaucoup plus nombreux » (Prost, 2001, p. 77). Lorsqu'en 1975 cette mesure est suivie par la création du « collège unique » – qui vise à « établir une meilleure égalité des chances devant l'enseignement » (p. 78) –, le système scolaire français doit « faire face au défi de la “massification” » (Meirieu & Le Bars, 2001, p. 113). Une « extension progressive à toute la population de scolarités initialement réservées à une minorité » – ou « démocratisation quantitative » de l'enseignement (Prost, 2001, p. 73) – semble effectivement avoir eu lieu ; cependant si l'on en croit la tendance lourde qui se dessine depuis vingt ans au gré d'études comparatives internationales comme PISA (OCDE⁷, 2001-2019), TIMSS⁸ (IEA⁹, 1995-2019) ou PIRLS¹⁰ (IEA, 2001-2021), la démocratisation *qualitative* de l'enseignement paraît encore aujourd'hui insuffisante en France.

1.2 Massification scolaire vs. démocratisation de la réussite

Tandis que selon (Toullec-Théry & Lescouarch, 2014 :10), « tout concourt à montrer que l'école française n'arrive pas à aider les élèves en grande difficulté », le constat de cet écart entre, d'une part, un progrès quantitatif de la scolarisation en France et, d'autre part, des difficultés rencontrées par l'École pour adjoindre à cette croissance des effectifs une hausse

⁷ OCDE : Organisation de coopération et de développement économiques.

⁸ TIMSS : « *Trends in International Mathematics and Science Study* ».

⁹ IEA : « *International Association for the Evaluation of Educational Achievement* ».

¹⁰ PIRLS : « *Progress in International Reading Literacy Study* ».

qualitative en termes de réussite pour les élèves, est également fait par (Duru-Bellat & Kieffer, 1999, p. 13), qui constatent une « absence de démocratisation qualitative [de l'enseignement] », due à une « stabilité des inégalités inhérentes au déroulement des carrières [scolaires], avec comme résultante une translation plus qu'une atténuation des inégalités ». Le paradigme de l'apprentissage tout au long de la vie pourrait constituer une forme de réponse à ces problématiques : en effet cette approche repose notamment sur le développement de la confiance en soi, de l'autonomie et de la coopération des élèves (Algan, Huillery, *et al.*, 2018 ; Lüftenegger *et al.*, 2012), et implique de ce fait une « remonétisation » de la formation scolaire, au moyen de compétences transverses et transférables, sources d'adaptabilité et de créativité pour les futurs professionnels que sont les apprenants (Demirel, 2009 ; Håkansson Lindqvist *et al.*, 2020).

En effet, « si la République a réussi, à force de réformes institutionnelles, à démocratiser l'accès à l'école, elle a échoué à démocratiser la réussite dans l'école » (Meirieu & Le Bars, 2001, pp. 111-112) car « si [les] méthodes [de l'école] ne forment pas à l'autonomie, à l'initiative, à la discussion critique, elles entérinent et renforcent les inégalités sociales » (p. 106). Or d'après le rapport PISA de 2018 (OCDE, 2019b, p. 1), le système scolaire français tend à creuser les inégalités sociales :

« Comme déjà observé lors des éditions précédentes du PISA, la France est l'un des pays de l'OCDE où le lien entre le statut socioéconomique et la performance dans PISA est le plus fort avec une différence de 107 points entre les élèves issus d'un milieu favorisé et ceux issus d'un milieu défavorisé. Cette différence est nettement supérieure à celle observée en moyenne dans les pays de l'OCDE (89 points). »

Si la massification scolaire représente un tel défi, c'est probablement parce qu'elle s'accompagne d'une « montée de l'hétérogénéité des élèves », notamment « dans les classes de collège » (Meirieu & le Bars, 2001, p. 26). La section suivante propose de définir cette notion d'hétérogénéité en contexte scolaire et les conditions sociohistoriques de son émergence en France, avant d'aborder certaines spécificités contextuelles, mais aussi structurelles, pouvant expliquer que le collège unique français soit particulièrement marqué par ces phénomènes.

1.3 Hétérogénéités contextuelles et structurelles du collège en France

Les hétérogénéités au collège sont liées aux phénomènes d'inégalités socioscolaires de diverses manières. Des effets contextuels sont susceptibles d'expliquer le renforcement des inégalités socioscolaires au collège, *e.g.* des différences inter-établissements. Précisément, la dominante en termes de catégories socioprofessionnelles des parents (collèges plutôt « populaires » vs. collèges plutôt « favorisés ») déterminerait un « contexte de scolarisation » pouvant produire des différences pour les élèves sur le plan de la progression, de la notation ou de l'orientation. Par ailleurs, en raison d'effets cumulatifs, le collège semble creuser les inégalités socioscolaires, davantage que l'école primaire : si pour un élève donné, des difficultés ou lacunes s'étaient accumulées lors de sa scolarisation dans l'enseignement primaire, elles ont tendance à s'accroître de façon non négligeable au collège, notamment en raison de biais sociocognitifs dans l'évaluation des acquis.

Un autre facteur, structurel cette fois, est susceptible d'expliquer partiellement l'accroissement des inégalités socioscolaires au collège : il s'agit de la constitution de classes de niveau. Cette pratique ne correspond pas aujourd'hui à une volonté délibérée de la part de l'institution scolaire : la constitution de classes de niveaux irait à l'encontre des tendances aujourd'hui observées dans les textes officiels de l'institution scolaire. Selon (Meirieu & Wagnon, 2018, p. 8), cela n'en demeure pas moins une réalité, notamment à travers le choix de spécialités et d'options, ou la constitution de « sections pour privilégiés » qui reviennent, *in fine*, à créer des classes de niveau. Ce phénomène nous renseigne en négatif sur les hétérogénéités du collège, en ce qu'il constitue une tentative de les contrecarrer de manière détournée. Ces classes de niveaux de facto témoignent de la prégnance des hétérogénéités dans le projet initial du collège unique ; cependant elles contribuent au renforcement des inégalités socioscolaires au collège.

Cette question des groupes de niveau est sous-tendue par une dichotomie entre classes homogènes et classes hétérogènes ; or en contexte hétérogène, comment permettre à tous les élèves de progresser ? C'est ici qu'intervient la notion de différenciation pédagogique, elle-même sous-tendue par l'idée d'hétérogénéité. Nous allons aborder ces questions dans les sections qui suivent.

1.4 L'hétérogénéité en milieu scolaire : définition et contextualisation

Cette hétérogénéité est une « notion très vaste » (Forget, 2017b), que Burns (1971, cité par Astolfi, 1995) tente d'opérationnaliser à travers sept postulats :

1. Il n'y a pas deux élèves qui progressent à la même vitesse ;
2. Il n'y a pas deux élèves qui soient prêts à apprendre en même temps ;
3. Il n'y a pas deux élèves qui utilisent les mêmes techniques d'étude ;
4. Il n'y a pas deux élèves qui résolvent les problèmes de la même manière ;
5. Il n'y a pas deux élèves qui possèdent le même répertoire de comportements ;
6. Il n'y a pas deux élèves qui possèdent le même profil d'intérêt ;
7. Il n'y a pas deux élèves qui soient motivés pour atteindre les mêmes buts.

C'est aussi une notion protéiforme, multidimensionnelle : (Meirieu & le Bars, 2001, p. 31) évoquent notamment une « hétérogénéité des niveaux, des motivations, des origines sociales, mais aussi des cultures et des systèmes de valeurs, des comportements scolaires et des projets professionnels des élèves ». À cela vient s'ajouter une hétérogénéité des difficultés en fonction du nombre d'élèves à profils spécifiques (*e.g.* dyslexie, dyspraxie, TDA/H¹¹, HPI¹²) ou du nombre d'élèves allophones de nationalités différentes.

Or les systèmes éducatifs qui ont opté pour une unification de structure (*e.g.* le « collège unique » en France) se sont trouvés ipso facto confrontés à une montée des hétérogénéités au sein des établissements et des classes. Ceci a nécessité un changement de paradigme, « d'une logique de différenciation/séparation, qui gère les difficultés d'apprentissage en séparant les publics d'élèves et en les orientant sur des voies différentes, vers une logique d'intégration qui implique une gestion plus directement pédagogique des difficultés d'apprentissage et de l'inévitable hétérogénéité des acquis, des rythmes et des aptitudes » (Lafontaine, 2017, p. 15).

La réponse institutionnelle la plus souvent mise en avant pour relever ce défi des hétérogénéités au collège est la différenciation pédagogique, qui fait l'objet du prochain chapitre.

¹¹ TDA/H : Troubles de l'attention avec ou sans hyperactivité.

¹² HPI : Haut potentiel intellectuel.

Ce qu'il faut retenir

En France, l'unification structurelle du système éducatif au moyen du « collège unique » s'est mécaniquement accompagnée d'une montée des hétérogénéités dans les classes. De ce fait, la massification scolaire n'a pas donné lieu à une démocratisation de la réussite et les inégalités socioscolaires ont eu tendance à s'accroître. En réponse à cela, une approche a été particulièrement mise en avant par l'institution scolaire : la différenciation pédagogique.

2 Vers une opérationnalisation de la différenciation pédagogique

Pour commencer nous allons nous intéresser aux origines de la différenciation pédagogique (DP), en tant qu'approche privilégiée par l'institution scolaire française pour répondre au défi des hétérogénéités. Nous proposerons ensuite un bref tour d'horizon de la DP, avant d'en souligner quelques limites et d'envisager un dépassement.

2.1 Une approche privilégiée par l'institution

La différenciation pédagogique constitue un point d'entrée essentiel, à la fois pour comprendre les enjeux socioscolaires liés à la prise en compte des hétérogénéités au collège unique, et pour la construction du cadre théorique de cette recherche. Dans le prolongement de (Legrand, 1995), elle est définie par (Forget, 2017b, p. 29) comme le fait, « à partir d'un objectif commun défini, [de] prendre en compte la diversité individuelle pour conduire tous les élèves à la maîtrise des objectifs fondamentaux ».

C'est en effet pour faire face au défi de l'hétérogénéité qu'à « partir des années quatre-vingts [...] la "pédagogie différenciée" se développe » (Meirieu & le Bars, 2001, pp. 25-26). En 1983, le ministre de l'Éducation nationale Alain Savary se déclarait ainsi en faveur d'un collège qui ne serait ni uniforme ni ségrégatif, mais diversifié. Dès 1985, les programmes et instructions pour l'école primaire et les collèges indiquent, pour chaque degré d'enseignement, que « le maître a recours à des pratiques pédagogiques différenciées, adaptées aux rythmes, aux difficultés et à la diversité des enfants » et qu'une « diversification et une individualisation de l'enseignement sont nécessaires pour répondre aux problèmes posés par les difficultés de certains élèves et l'hétérogénéité des classes » (Colas Degenne, 2016, p. 28). Cette tendance se

confirme par la suite : dans la Loi d'orientation sur l'Éducation de 1989, on peut par exemple lire que « les enseignants organisent leurs enseignements en prenant en compte les rythmes d'apprentissage de chaque élève » (République française, 1989, article 5). Plus récemment, la différenciation pédagogique figure parmi les axes prioritaires de la formation continue pour l'année scolaire 2017-2018 (Ministère de l'Éducation nationale, 2017). De même, il est stipulé dans le Code de l'éducation (République française, 2000, article L111-1) que celle-ci est la « première priorité nationale » et qu'elle « contribue à l'égalité des chances et à lutter contre les inégalités sociales et territoriales en matière de réussite scolaire et éducative », notamment en veillant « à la scolarisation inclusive de tous les enfants, sans aucune distinction » et « à la mixité sociale des publics scolarisés au sein des établissements d'enseignement ».

2.2 Une approche difficile à opérationnaliser

Malgré ces prescriptions institutionnelles déjà bien ancrées, la différenciation pédagogique paraît difficile à opérationnaliser, notamment dans la recherche. Dès 1997, Duru-Bellat & Mingat signalaient en effet que :

« L'idée qu'il convient de traiter différemment les élèves, si l'on entend leur faire atteindre des objectifs communs, est devenue centrale dans les conceptions « pédagogiquement correctes » du moment. Les acteurs [du monde de l'éducation] se sont emparés de la justification de la notion de pédagogie différenciée, sous-jacente au thème de « l'indifférence aux différences » [Bourdieu & Passeron, 1970] et ont développé la notion de « discrimination positive ». Mais la question des effets de ces traitements pédagogiques différenciés, sur les acquisitions ou les attitudes des élèves, est restée, en France, très peu documentée » (Duru-Bellat & Mingat, 1997, p. 760).

Cette situation semble avoir évolué depuis, tout du moins en ce qui concerne la recherche fondamentale. A cet égard (Forget, 2017b) a montré que la différenciation pédagogique constitue un champ plutôt bien investi compte tenu du grand nombre de modèles d'action qu'on peut recenser. Pour autant, « non seulement les propositions [de différenciation pédagogique] se déclinent sous des formes multiples mais chacune d'elles met de surcroît en scène un nombre impressionnant de variables de fonctionnement formant un tout complexe

difficilement contrôlable par un plan de recherche expérimentale ». Il s'ensuit une difficulté à établir par des études empiriques des préconisations opérationnalisables en termes de DP.

En France la réalité concrète des pratiques de terrain semble d'ailleurs s'en ressentir, en particulier au collège. A ce titre, dans l'enquête internationale TALIS-collège menée en 2013, seulement 22 % des enseignants français déclarent qu'ils « donnent des travaux différents aux élèves qui ont des difficultés d'apprentissage et/ou à ceux qui peuvent progresser plus vite », alors que pour l'ensemble des pays de l'OCDE participant à l'enquête, la moyenne se situe à 44 % (données de l'enquête internationale TALIS 2013 [OCDE, 2014], reprises par [Cnesco, 2017] – voir également la Figure 1).

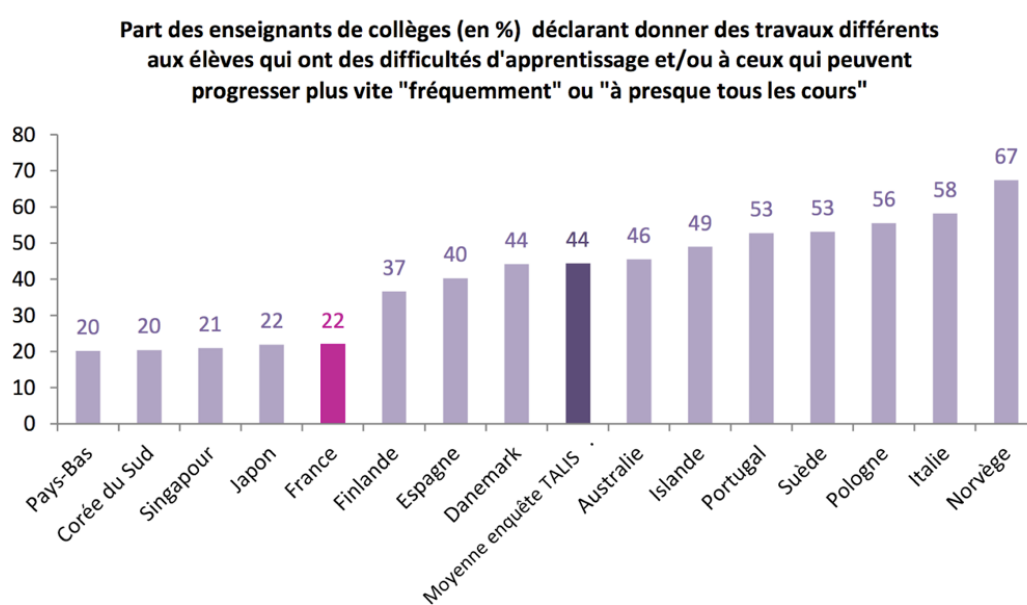


Figure 1 – Pratiques déclarées de différenciation pédagogique par les enseignants des pays de l'OCDE : données de l'enquête internationale TALIS 2013 (OCDE, 2014), reprises par (Cnesco, 2017)

Ce constat concernant les pratiques déclarées par les enseignants français sur le plan de la différenciation pédagogique fait écho aux résultats de l'enquête PISA 2018 (OCDE, 2019b, p. 1), notamment eu égard au soutien relativement faible ressenti par les apprenants français en provenance de leurs enseignants :

« La France est l'un des pays où les élèves ressentent le moins de soutien de la part de leurs enseignants pour progresser dans les apprentissages. Ainsi, moins d'un élève sur quatre en France (un sur trois, en moyenne dans les pays de l'OCDE) déclare que son professeur lui indique ses points forts. De même, moins de deux élèves sur cinq en France, alors que près d'un élève sur deux en moyenne dans les pays de l'OCDE, déclarent

que leur professeur leur indique souvent ou toujours comment améliorer leurs résultats. »

Enfin, un troisième résultat de l'enquête PISA 2018 peut être mis en lien avec la problématique de la différenciation pédagogique : en France, « seuls 45 % des élèves ont déclaré que les élèves coopèrent entre eux dans leur établissement (moyenne OCDE : 62 %) » (OCDE, 2019, p. 9).

Malgré tout, on peut recenser un certain nombre de pratiques, de dispositifs et de gestes professionnels en lien avec la différenciation pédagogique, et qui font l'objet de la section suivante.

2.3 La différenciation pédagogique en pratique

La différenciation pédagogique peut être envisagée comme un continuum : aux extrémités du spectre, il y aurait la différenciation structurelle d'une part, qui se joue à l'extérieur de la classe, et la différenciation pédagogique stricto sensu d'autre part, qui se joue en interne, à l'intérieur même de la salle de classe (Forget, 2017a; Lafontaine, 2017).

2.3.1 A l'extérieur de la classe : la différenciation structurelle

Au départ les élèves sont essentiellement regroupés en fonction de leur âge, dans des classes hétérogènes en termes socioculturels et de maîtrise des compétences. La différenciation structurelle consiste alors à créer de manière permanente d'autres groupes, moins hétérogènes, ayant vocation à accueillir et regrouper des élèves différents mais qui se ressemblent entre eux. Cette approche entre alors en tension avec la tendance hétérogène du mode de regroupement habituel. En France, les classes de niveau n'existent pas officiellement ; cependant au collège certaines pratiques aboutissent, de fait, à la création de « classes de niveau qui ne disent pas leur nom » (Meirieu & Wagnon, 2018, p. 8), notamment grâce à des spécialités et filières en lien avec un projet sportif, artistique ou professionnel, ou par le choix d'options spécifiques (Lafontaine, 2017, p.10), « dont la difficulté relative ou la rareté engendre de facto des hiérarchies et donc des regroupements par aptitude (par exemple : enseignement immersif [*i.e.* « classes bilingues »], choix des langues anciennes ou du chinois) ».

D'autre part, dans le système scolaire français, des réformes ont été mises en œuvre dès 1989 contre le redoublement – qui consiste à regrouper des élèves ayant des difficultés importantes avec d'autres élèves moins âgés, donc moins avancés sur le plan de la maîtrise des

compétences. Le redoublement est ainsi devenu très rare en France (Kahn, 2010, p. 74). Toutefois la carte scolaire continue à être une source de différenciation structurelle/externe en France. Puisqu'elle « oblige les enfants à s'inscrire dans l'école de leur quartier », elle sépare de fait « les élèves des quartiers difficiles de ceux des centres-villes » et génère des disparités socioéconomiques : elle « s'applique de façon inégalitaire aux différents groupes sociaux » (Oberti, 2007, p. 262). Elle est de surcroît « contournée par ceux qui ont les moyens de bénéficier de passe-droits ou de financer des études dans un établissement privé » (Attali, 2008, p. 28). Cette source de différenciation pédagogique structurelle implique un cumul des dimensions spatiales et sociales qui « renforce la protection des plus favorisés, fragilise les situations de mixité dans les collèges « ordinaires », et enfin accentue la relégation et la disqualification des plus défavorisés » (Oberti, 2007, p. 262).

2.3.2 A l'intérieur de la classe : la différenciation pédagogique *stricto sensu*

La différenciation peut aussi s'opérer en interne : on peut alors parler de différenciation pédagogique *stricto sensu*. Cette approche consiste avant tout à mettre en œuvre un « cadre souple où les apprentissages sont suffisamment explicités et diversifiés pour que les élèves puissent travailler selon leurs propres itinéraires d'appropriation tout en restant dans une démarche collective d'enseignement des savoirs et savoir-faire communs exigés » (Forget, 2017b, p. 17). Cependant, vu la pluralité des démarches existantes, « le caractère hautement polysémique du lexique » désignant les concepts et la variété de thématiques convoquées par ce champ, on parlera volontiers de « pédagogies différenciées » au pluriel (Forget, 2017b, p. 17).

D'après (Forget, 2017a), « la plupart des auteurs sont néanmoins unanimes sur l'idée que la DP peut toucher les contenus, les processus et les productions ». Dans le prolongement de cette typologie, (Forget & Lehraus, 2015) proposent de distinguer trois temporalités : avant, pendant et après la situation d'enseignement/apprentissage. Cette modélisation permet de distinguer différents « gestes de différenciation », tels que : tester (évaluation diagnostique), réactiver, préparer, soutenir, adapter, évaluer (au sens formatif du terme), exercer, revoir, évaluer (au sens certificatif du terme). La Figure 2 récapitule cette modélisation de la DP selon (Forget & Lehraus, 2015).

<p><u>Différencier AVANT l'enseignement d'une notion : dans quels buts?</u></p> <ol style="list-style-type: none"> 1) Tester : avant de mettre un objet d'enseignement à l'étude, il peut être utile de tester les prérequis de certains élèves pour lesquels on ne situe pas précisément le niveau. 2) Réactiver : le fait de réactiver les notions utiles pour l'enseignement qui va suivre peut constituer une aide efficace pour certains. En effet, tous les élèves ne sont pas égaux en matière de conscience disciplinaire. 3) Préparer : l'enseignant peut préparer certains élèves à une activité en leur fournissant des clés d'accès, des repères avant l'activité : ceci permet aux élèves concernés de suivre ensuite avec moins de difficultés l'activité avec le groupe. <p><u>Différencier PENDANT l'enseignement d'une notion : dans quels buts?</u></p> <ol style="list-style-type: none"> 1) Soutenir : l'enseignant peut soutenir ou étayer les élèves de manière plus ou moins rapprochée eu égard à leur niveau de compétence. 2) Adapter : l'adaptation des conditions de réalisation d'une tâche englobe à la fois l'aménagement des supports (matériel), de la consigne, des exigences, de la quantité de travail ou encore du temps mis à disposition. 3) Évaluer (formatif) : nous n'avons cependant pas besoin à tout moment de recueillir de l'information sur chaque élève. Le recueil est parfois ciblé sur les élèves dont la progression pose précisément problème. <p><u>Différencier APRÈS l'enseignement d'une notion : dans quels buts?</u></p> <ol style="list-style-type: none"> 1) Exercer : tous les élèves ne sont pas égaux face à l'automatisation des procédures apprises : certains seront prêts avant les autres, d'autres auront besoin de plus ou moins de temps pour une consolidation. 2) Revoir : les difficultés d'un élève ne sont pas celles d'un autre. La révision de certains éléments non acquis sera ciblée sur les difficultés réelles des élèves. 3) Évaluer (certificatif) : certains élèves ont besoin que les conditions de réalisation et/ou critères des évaluations soient adaptés à leurs difficultés particulières (par exemple, élèves dyslexiques, allophones).

Figure 2 – Modélisation de la DP par (Forget & Lehraus, 2015, p. 75)

Une autre manière d'analyser les pratiques de différenciation pédagogique dans la classe est celle de (Meirieu, 2015, pp. 134-135). Cet auteur distingue entre différenciation successive et différenciation simultanée. La différenciation successive consiste à diversifier les approches pédagogiques au sein du cours, *i.e.* à utiliser pendant le cours une variété d'outils et de situations d'apprentissages, afin de parvenir à l'objectif didactique que l'enseignant s'est fixé. La diversification pédagogique qui en résulte permet à priori de mieux répondre aux besoins d'élèves aux profils variés, dans un contexte marqué par l'hétérogénéité des groupes classes. La différenciation simultanée implique une décentration de l'enseignant et consiste, quant à elle, à faire coexister dans l'espace même de la classe, en même temps, des modalités d'enseignement/apprentissage diverses, *e.g.* en organisant différents ateliers – à l'instar des pratiques qu'on peut observer dans des écoles Montessori (Gardes & Courtier, 2018).

Pour organiser la différenciation simultanée telle que (Meirieu, 2015) la définit, on pourra envisager plusieurs dispositifs pédagogiques favorisant la DP, tels que la « table d'appui », qui permet notamment un étayage par l'adulte pour venir en aide aux élèves les plus en difficultés, regroupés pour la circonstance autour d'une même table, pour une remédiation ponctuelle et ciblée. Pour autant ce dispositif permet également de favoriser l'autonomie des élèves en abordant des aspects méthodologiques et en contribuant au développement de la métacognition des apprenants (Bruner, 1983). De même dans le prolongement de (Crahay & Delhaxhe, 2003), on peut envisager la création de groupes de besoins homogènes : des

groupements « flexibles et [dont la] pertinence est régulièrement réévaluée eu égard aux progrès des élèves » cité par (Forget, 2017a, p. 23).

2.3.3 Quelques dispositifs clés

On peut enfin recenser plusieurs dispositifs couramment employés pour la différenciation pédagogique tels que le plan de travail, la table d'appui ou les classes flexibles. Le plan de travail permet d'organiser le travail de l'élève en lui fixant des objectifs d'apprentissage, tout en lui laissant une marge de manœuvre, notamment quant à l'ordre dans lequel il va les atteindre. Il s'agit également de distinguer entre des compétences de base, dont la maîtrise est indispensable – *i.e.* celles du Socle commun de connaissances, de compétences et de culture –, et des compétences subsidiaires d'approfondissement ou de perfectionnement (Forget, 2017b). Lorsque les élèves travaillent en autonomie, la table d'appui est un espace dédié à l'accueil d'élèves ou de petits groupes de niveaux que l'enseignant choisira d'accompagner ponctuellement en fonction des besoins. Les classes flexibles consistent en un aménagement varié de l'espace classe (*e.g.* îlots, disposition en *U*, en « rangs d'oignon », coin lecture...), au moyen d'un mobilier ad hoc (*e.g.* assises stimulantes ou calmantes, tables de différentes hauteurs, rangements divers, cloisons flexibles...) de telle sorte que les élèves puissent opter pour la modalité d'apprentissage qui leur convient. Il s'agit de favoriser un engagement actif dans les apprentissages, avec une autonomie et une responsabilisation grandissante, ainsi qu'une conscientisation accrue en termes de besoins et de réussite, avec un impact positif sur la maîtrise des compétences (Hoffman, 2002; Miller-Cochran & Gierdowski, 2013; Rytivaara, 2011). Les classes flexibles s'accompagnent volontiers d'environnements informatisés pour l'apprentissage humain (EIAH), notamment pour des apprentissages hybrides (Müller & Mildenerger, 2021), ainsi que de modalités d'interactions coopératives entre élèves, comme le tutorat, l'aide, l'entraide et le travail de groupe (Connac, 2022).

Ceci dit la différenciation pédagogique se heurte à des difficultés de mise en œuvre et à des limites intrinsèques que nous allons essayer d'explicitier maintenant.

2.4 Limites de la différenciation pédagogique

Nous avons montré que, dans le prolongement de la création du collège unique en 1975, la différenciation pédagogique (DP) s'était rapidement imposée en tant que réponse institutionnelle privilégiée pour relever le défi des hétérogénéités au collège. Nous avons également vu qu'en réalité la DP correspond à une pluralité d'approches, structurées autour

d'une tension entre la différenciation institutionnelle d'une part, et la différenciation pédagogique au sens strict d'autre part. Concrètement, la DP se traduit par ailleurs par des dispositifs de soutien externalisés (Houssaye, 2012). Ce soutien externalisé peut s'exercer individuellement (*e.g.* tutorat enseignant-élève) ou en groupe (*e.g.* ULIS¹³, « Devoirs faits », demi-groupes du dispositif « Aide personnalisée »). Cette « externalisation des élèves » (Connac, 2016, p.1) ne favorise pas toujours l'inclusion scolaire des élèves en difficulté. Plusieurs études récentes (Cnesco, 2017; Connac, 2018; Kahn, 2015) soulignent ainsi les limites de ces approches : malaise scolaire, stigmatisation, répétition et intériorisation des situations d'échec personnel) – autant d'éléments susceptibles de nuire aux apprentissages (Lescouarch, 2021 : 8).

Outre ces difficultés liées à l'externalisation des dispositifs de différenciation – externalisation qui, au final, tend à renforcer les inégalités socioscolaires qu'elle prétend combattre –, on constate l'existence, sur le terrain, d'un « effet élastique » (Laffitte *et al.*, 2006). Il s'agit en quelque sorte d'un mouvement de balancier par lequel des enseignants déterminés, pleins de bonne volonté, mettent en œuvre dans leur classe un certain nombre de dispositifs innovants, qui autonomisent par exemple les élèves et favorisent la coopération ; malheureusement ces approches pédagogiques s'avèrent trop difficiles à réguler, par exemple en regard de leurs compétences professionnelles du moment. Le résultat final est contreproductif (indiscipline, « chahut », retards dans les apprentissages) ce qui conduit l'enseignant à revenir à des méthodes plus traditionnelles – *i.e.* le « cours dialogué » ou « semi-magistral » (Connac, 2022, p. 11).

Par ailleurs, la différenciation successive et la différenciation simultanée selon (Meirieu, 2015), évoquées plus haut, génèrent des difficultés dans la classe. D'une part, la différenciation successive peut s'avérer trop chronophage, à court-terme eu égard aux contraintes organisationnelles du collège – avec ses cours dont la durée est généralement comprise entre 55 minutes et 1h55–, mais aussi sur une temporalité plus longue, en raison du schéma directeur le plus couramment observé pour la mise en œuvre des séquences d'enseignement/apprentissage : « leçons, entraînements, révisions, évaluations » (Connac, 2022, p. 22). D'autre part, la différenciation simultanée implique de la part de l'enseignant un effort de préparation significatif en amont des séances avec les élèves. Pour une séance donnée, si l'on compare avec un modèle plus « traditionnel » – *e.g.* en cours dialogué –, la

¹³ ULIS : Unité locale d'inclusion scolaire.

différenciation simultanée implique en quelque sorte une démultiplication des supports et des scénarios pédagogiques.

Dans ce contexte, « trente ans après la loi de 1989 généralisant en France le principe d'une différenciation pédagogique sans que les pratiques ne se généralisent » (Lescouarch, 2021 : 51), les pédagogies coopératives pourraient représenter une alternative viable, fondée notamment sur la notion de collectif apprenant et sur des étayages variés, implémentés de manière collégiale.

Ce qu'il faut retenir

Dès 1985, la différenciation pédagogique est mise en avant par l'institution scolaire pour répondre au défi des hétérogénéités au collège. Bien que la recherche fondamentale ait produit un nombre relativement élevé de modèles d'action en la matière, ces derniers s'avèrent difficiles à opérationnaliser dans le cadre d'études empiriques. En outre, les pratiques déclarées des enseignants montrent que la différenciation pédagogique s'avère difficile à mettre en œuvre concrètement sur le terrain. Qu'il s'agisse de différenciation structurelle ou strictement pédagogique, en appui sur certains dispositifs clés comme les classes flexibles, la différenciation pédagogique se heurte à des limites que les pédagogies coopératives pourraient permettre de dépasser.

3 Les pédagogies coopératives

Pour commencer, nous tenterons une désambiguïsation des termes *coopération* et *collaboration*, qui semblent à la fois synonymes et concurrents. Nous examinerons alors ces deux paradigmes voisins mais distincts, afin d'envisager des pistes pour l'opérationnalisation de la différenciation pédagogique au moyen des pédagogies coopératives. Nous verrons notamment que les notions de coopération et d'apprentissages collaboratifs occupent une place significative dans les curriculums et les textes officiels du ministère de l'Éducation nationale.

3.1 Coopération vs. collaboration : tentative de désambiguïsation

Les termes « coopération » et « collaboration » semblent utilisés indifféremment, y compris dans les textes officiels. Nous nous proposons d'examiner maintenant l'étymologie

puis la littérature scientifique, pour voir s'il est possible d'entrevoir des distinctions entre les deux mots.

Pour commencer, au sens étymologique, « co-opérer » veut dire « opérer ensemble ». Le Trésor de la langue française précise que ce terme signifie « agir, travailler conjointement avec quelqu'un en vue de quelque chose, participer, concourir à une œuvre ou à une action commune ». De même, pour le Larousse, coopérer c'est « prendre part, concourir à une œuvre commune ; contribuer, participer ». En ce qui concerne le mot « collaboration » (du bas latin « *collaborare* », travailler avec quelqu'un), la 9^e édition du dictionnaire de l'Académie française indique qu'il s'agit du « fait de travailler avec une ou plusieurs personnes à une œuvre commune » ; de même, pour le Larousse, collaborer signifie « travailler de concert avec quelqu'un d'autre, l'aider dans ses fonctions ; participer avec un ou plusieurs autres à une œuvre commune ». Quelles que soient les sources, ces deux termes semblent donc très proches, potentiellement interchangeables. Dans la littérature scientifique, la coopération et la collaboration sont des objets de recherche travaillés dans une variété de disciplines (Thomson *et al.*, 2008, p. 97) et il ne semble pas y avoir de consensus terminologique en ce qui concerne ces deux termes (Sanojca, 2018, p. 37; Thomson & Perry, 2006, p. 23). Dans leurs travaux, des auteurs de différents champs disciplinaires ont ainsi proposé plusieurs approches qui pourraient nous aider à mieux distinguer ces deux mots et à limiter les risques d'ambiguïté.

Chez des chercheurs en management, on trouve par exemple la définition suivante :

« La collaboration est un processus par lequel des acteurs autonomes interagissent à travers des négociations formelles et informelles, créant conjointement les règles et structures gouvernant leurs relations et des moyens d'agir ou de prendre des décisions concernant les problèmes qui les ont réunis ; c'est un processus impliquant des normes partagées et des interactions mutuellement bénéfiques » (Thomson & Perry, 2006, p. 23, traduction libre).

Selon ces auteurs, cette définition de la collaboration implique « un degré plus élevé d'action collective que la coopération ou la coordination » (Thomson & Perry, 2006, p. 23, traduction libre). Dans le champ des sciences de l'éducation et de la formation, d'autres auteurs proposent cette distinction « hiérarchique » entre coopération et collaboration, à l'instar de (Baker, 2008, p. 8), selon qui « l'apprentissage coopératif » serait « la dénomination de tout type d'apprentissage produit dans une situation de travail de groupe, et « l'apprentissage collaboratif » désignerait l'apprentissage produit grâce à une véritable collaboration ». De

même, avec (Henri & Lundgren-Cayrol, 2001), les apprentissages coopératifs sont considérés comme mieux adaptés à des apprenants plus jeunes car ils requièrent moins d'autonomie et de maturité, et permettent à l'enseignant d'exercer un certain contrôle (« hétéro-direction » des apprentissages). Au contraire, les apprentissages collaboratifs seraient davantage conseillés pour des apprenants adultes car ils nécessitent une autonomie et une maturité plus grandes, en raison de la marge d'autodirection plus importante qui leur est laissée. Pour autant, des auteurs comme (Cavalier *et al.*, 1995; Henri & Lundgren-Cayrol, 2001) s'accordent à dire que les apprentissages coopératifs constituent en quelque sorte une initiation aux apprentissages collaboratifs. Selon ces auteurs, coopération et collaboration participent d'un même principe éducationnel et s'inscrivent dans un continuum.

Alors qu'il paraît théoriquement difficile de désambigüiser ces deux termes, (Connac, 2017b, p. 20) introduit une distinction d'ordre praxéologique entre coopération et collaboration. Pour lui, la collaboration renvoie étymologiquement au travail (du latin « *labor* »), en vue d'un but mutuellement bénéfique, et donc avec « un partage d'intérêts qui peut être à la fois solidaire et égoïste ». Selon cette conception, le travail collaboratif se traduit par des travaux individuels coordonnés, *i.e.* du « travail de groupe ». A contrario, la coopération se concrétise par des relations d'aide, de tutorat et d'entraide, ainsi que par du « travail *en* groupe », *i.e.* une forme de travail où l'activité de chacun est fortement interdépendante de celle d'autrui. Ainsi envisagée, la coopération se situerait davantage dans une perspective de gratuité altruiste (Go, 2013). La Figure 3, issue de (Connac, 2017b, p. 20), schématise ce distinguo.

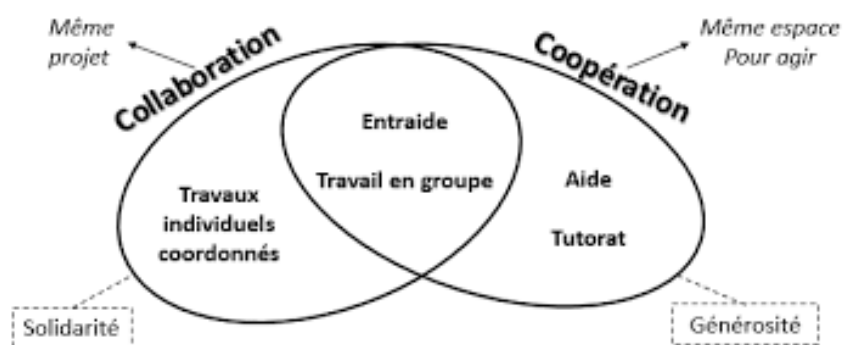


Figure 3 – Coopération versus collaboration selon (Connac, 2017b, p. 20)

Pour ce travail de thèse, face à la variabilité terminologique constatée en fonction des auteurs, et dans la mesure où, au plan épistémologique, notre recherche-intervention est collaborative, nous avons souhaité éviter tout risque de confusion et avons opté pour la convention suivante : nous utiliserons le terme *collaboration* lorsqu'il sera question de la recherche collaborative entre le chercheurs, les enseignants partenaires et les autres acteurs

concernés ; tandis que lorsqu'il s'agira des apprentissages coopératifs des élèves, amenés à aider leurs pairs, à s'entraider, à accomplir un travail *de* groupe, un travail *en* groupe – voire, à certains moments, à coopérer avec l'enseignant (*e.g.* discussions de classe) –, nous emploierons le terme *coopération*.

3.2 Origines et place dans les textes officiels

L'ancrage historique des apprentissages coopératifs est relativement ancien ; on en retrouve notamment des traces dans l'antiquité. A ce titre, on peut citer le philosophe romain Sénèque qui, au I^{er} siècle de l'ère commune, faisait déjà indirectement l'apologie des apprentissages coopératifs à travers des affirmations comme « *Homines dum docent discunt* » (« Les hommes apprennent en enseignant »). De même, dans le Judaïsme, dès le II^e siècle de l'ère commune, l'étude talmudique « s'accomplit généralement au sein de paires d'apprenants », selon le principe de la « *chavruta* » ou étude dyadique – une pratique qui continue encore aujourd'hui dans certains milieux juifs orthodoxes (Schwarz, 2020, p. 3). Au XVII^e siècle, le pédagogue européen d'origine tchèque Jan Amos Komenský (« Comenius », 1592-1670) soulignait lui aussi le caractère mutuellement bénéfique de la coopération entre pairs, considérant que « les apprenants gagneraient à enseigner aux autres élèves, autant qu'ils gagneraient à être enseignés par eux » (Johnson & Johnson, 2017, p. 2, traduction libre). Si le paradigme du *cooperative learning* a été théorisé aux Etats-Unis dans les années 1970 (Reverdy, 2016, p. 2), les approches coopératives étaient déjà présentes par ailleurs, du XVIII^e siècle à nos jours – un nombre significatif de pédagogues ayant contribué à ce champ. Parmi les précurseurs, on pourra notamment citer Pestalozzi (1746-1827), Lancaster (1778-1838), Dewey (1859-1952), Montessori (1870-1952), Claparède (1873-1940) ou Cousinet (1881-1973). En France, le courant des pédagogies actives et le mouvement de l'Education nouvelle ont fortement contribué à l'essor des pédagogies coopératives, notamment grâce aux travaux de Célestin et Élise Freinet ou Fernand Oury.

Par ailleurs, tout comme la différenciation pédagogique, les pédagogies coopératives prennent rapidement une place importante dans les textes officiels. Notamment, dans le prolongement de la Loi d'orientation de 1989 sur l'Education, le Code de l'Education stipule au début des années 2000 que « le service public de l'éducation [...] favorise la coopération entre les élèves » (République française, 2000, article L111-1). Cette tendance se poursuit dans les années qui suivent, puisque les Programmes officiels actuels, qui sont entrés en vigueur en 2016 et ont été « consolidés » en 2018 et 2020, ont occasionné une refonte du Socle commun

de connaissances, de compétences et de culture, dont « le domaine 2 (...) concerne l'apprentissage du travail coopératif et collaboratif sous toutes ses formes » (Education nationale, 2018, p. 5).

Plus spécifiquement, pour ce qui à trait à l'enseignement/apprentissage du français en général, et de la production d'écrit en particulier, les Programmes actuellement en vigueur précisent que « le regard positif de l'enseignant qui encourage l'élève, le caractère motivant des différentes situations proposées, porteuses de sens, la *collaboration entre pairs* donnent le plaisir de l'écriture et aiguisent la curiosité pour la langue et son fonctionnement. » (Education nationale, 2018, p. 17).

3.3 Approches contemporaines

Parmi les chercheurs contemporains qui ont investi le champ des pédagogies coopératives, on peut citer notamment (D. W. Johnson & Johnson, 2002, 2009, 2017), ou (Lescouarch, 2010, 2020). La mise en œuvre des pédagogies coopératives s'avère fructueuse en termes d'apprentissages, de développement de stratégies d'entraide et d'épanouissement socioaffectif pour les apprenants (Ajaja & Eravwoke, 2010; Gillies, 2004).

Avec des auteurs comme (Reynaud, 2022) ou (Connac & Cieutat, 2021), une typologie émerge, qui propose de distinguer quatre formes de relations coopératives : l'aide, l'entraide, le travail en groupe et le tutorat. (Lescouarch, 2021) montre que ces modalités coopératives gagnent à s'inscrire dans une dynamique d'établissement ou à donner lieu à des collaborations resserrées au sein de l'équipe enseignante, pour fonder ce qu'il nomme un *collectif apprenant*.

Nous verrons ici que l'ensemble de ces modalités ou approches convoquent, chacune à leur manière et à des degrés divers, les notions de motivation, de métacognition et d'auto-efficacité – autrement dit, ce que (Carré, 2010) nomme *l'autodirection des apprentissages*, et qui fera l'objet du prochain chapitre.

3.3.1 L'aide

La relation d'aide est une interaction dissymétrique entre deux apprenants à un instant t : l'aidant manifeste une meilleure maîtrise de la compétence cible ou du faisceau de compétences mobilisées pour compléter un exercice, résoudre une tâche complexe ou une situation-problème (Barth, 2013) et se positionne ponctuellement en tant qu'expert vis-à-vis du pair qu'il aide (l'aidé). (Connac, 2017b, p. 24) précise que la verticalité induite par ce type d'interaction ne pose pas de problème, d'abord parce qu'elle s'effectue à la demande de l'aidé

lui-même, ensuite parce qu'elle implique une réversibilité possible, l'aidé pouvant devenir aidant à son tour – par exemple pour résoudre ultérieurement une situation-problème mobilisant d'autres compétences vis-à-vis desquelles il se positionne en tant que « pair-expert ».

Dans le prolongement des travaux de (Barth, 2013), le fait que l'aidé adopte une posture d'élève chercheur implique de privilégier ce que (Crahay *et al.*, 2001, p. 135) appellent des « tactiques rétroactives » d'aide, *i.e.* lorsque l'aidé a d'abord cherché à résoudre la tâche par lui-même et que l'aidant lui fournit une explication à posteriori, en « réagissant aux erreurs de [l'aidé] et/ou à ses questions ». Il s'agit d'une approche socioconstructiviste *rétroactive* de la relation d'aide. En effet, à contrario, une « approche proactive » risquerait de conduire à ce que l'aidé soit « dessaisi du travail cognitif d'élaboration de la stratégie » (Connac, 2017b, p. 24).

En s'appuyant notamment sur les travaux de (Baudrit, 2007), (Connac, 2017b, p. 25) distingue également trois degrés d'élaboration de l'aide fournie par l'aidant à l'aidé : l'aide exécutive est improductive car elle se borne à fournir immédiatement la réponse demandée, sans explication – cette modalité pouvant s'apparenter à une aide proactive ; l'aide peu élaborée fournit quant à elle une réponse élémentaire mais appropriée à une tâche peu compliquée, moyennant de bonnes compétences relationnelles de la part de l'aidant ; enfin l'aide élaborée semble la plus productive en termes socioconstructivistes dans la mesure où les explications fournies sont accompagnées d'un partage par l'aidant de *stratégies* pour résoudre la situation-problème.

L'aide élaborée implique donc une maîtrise avancée de la (ou des) compétence(s) travaillée(s), mais aussi des connaissances métacognitives (Berger & Büchel, 2012; Grangeat, 1999) à la fois déclaratives, procédurales et conditionnelles (Paris *et al.*, 1983), ainsi que le métalangage –ou « langage de second ordre » (VandenBos, 2007, p. 645) – nécessaire pour les partager avec l'aidé. En raison des compétences, des métaconnaissances et du métalangage mobilisés pour fournir une aide élaborée, on peut parler ici pour l'aidant de « métacompétence », *i.e.* « des qualités et capacités d'ordre supérieur et de nature conceptuelle, [personnelle et interpersonnelle] » (Bogo *et al.*, 2013, p. 260). Lorsque l'aidant fournit à l'aidé une telle aide, la dissymétrie initiale de la relation tend à se renforcer, car l'aidant voit ses compétences et métacompétences consolidées (Connac, 2017b, p. 26, 2022, p. 44). Cette aide élaborée comporte cependant un risque de « surétayage » qui renforcerait l'assistanat chez l'aidé : « placé en position de récepteur d'informations souvent trop étayantes, il ne pourrait entrer dans une démarche de résolution de problème et se contenterait d'exécuter des consignes » (Connac, 2017b, p. 26).

Avec (Baudrit, 2007, cité par Connac, 2017b), on peut alors caractériser une aide efficace au moyen de cinq critères : pertinence, relative brièveté du délai d'attente entre la demande d'aide et l'aide apportée, compréhension par l'aidé des explications fournies par l'aidant, résolution par l'aidé de la tâche qui lui posait problème et réinvestissement des explications fournies dans d'autres contextes similaires, degré d'élaboration de l'aide (exécutive, peu élaborée ou élaborée) adapté aux besoins de l'aidé. (Webb, 1989) propose une sixième condition : que l'aide soit à l'initiative de l'aidé – ce qui rejoint l'approche socioconstructiviste de (Barth, 2013) et la préférence de (Crahay *et al.*, 2001) pour l'aide rétroactive plutôt que proactive.

Ainsi, en s'appuyant sur les conclusions de (Berger & Büchel, 2012, p. 119 ; Schunk & Ertmer, 2000), le « modelage » entre pairs résultant d'une aide « ajustée », et l'appropriation de nouvelles stratégies métacognitives qu'elle induit potentiellement pour l'aidé, semblent de nature à développer ses « croyances motivationnelles » et son sentiment d'efficacité personnelle.

3.3.2 L'entraide

« L'entraide représente une interaction entre plusieurs personnes, conjointement bloquées face à une même difficulté. Ces personnes se réunissent de manière informelle, c'est-à-dire à leur initiative, pour tenter de la résoudre ensemble. Elles s'organisent librement » (Connac, 2017b, p. 27).

Ici la relation implique une symétrie, avec un niveau de maîtrise équivalent de la compétence cible – *i.e.* les binômes ou groupes d'entraide sont homogènes par nature (Baudrit, 2007). Cependant cette modalité coopérative ne stigmatise pas l'élève, puisqu'à la différence des groupes de niveaux usuels, c'est lui qui est ici à l'initiative. En effet, en vue d'un objectif d'apprentissage donné (tâche complexe, exposé, recherche...), l'entraide peut survenir de manière spontanée, à l'initiative d'un élève ; suite à quoi un petit groupe de travail peut se constituer sans intervention du professeur. Alternativement, lorsque des élèves sont confrontés à une même difficulté, dans le cadre d'une activité donnée, ils peuvent choisir de constituer un tel groupe d'entraide.

3.3.3 Le tutorat

Le tutorat consiste quant à lui en une relation dissymétrique impliquant un tuteur et un tutoré. Les deux membres de ce binôme bénéficient du dispositif, quoique différemment : le

tutoré est cognitivement placé en réception active, avec une composante métacognitive de transposition de savoirs, savoir-faire ou savoir-être, qu'il s'agit de s'approprier. Le tuteur est quant à lui placé d'emblée dans une dynamique métacognitive d'écoute ou de repérage de ce qui fait problème, et de didactisation de ses compétences pour en faciliter le transfert à destination du pair tutoré : pour comprendre ses erreurs en profondeur, pour se mettre à son niveau, il s'agit de monter en métacompétences. Le tutorat est destiné à approfondir ou à réviser des notions déjà introduites par ailleurs. Il doit aussi être scénarisé en amont de manière appropriée et s'appuyer sur des binômes compatibles, à la fois en termes d'aptitudes et d'attitudes. Dès lors, ce type d'interaction coopérative peut s'avérer porteur à la fois pour les apprentissages du tutoré et pour ceux du tuteur, et ce sans entraîner d'accroissement des inégalités socioscolaires dans la mesure où, au-delà d'un certain seuil, la marge de progression du tuteur devient minime (Crahay, 2013; Forget, 2017b).

3.3.4 Le travail en groupe

Le travail en groupe consiste à accomplir une tâche en étroite collaboration, en vue d'un but commun. Cette modalité coopérative nécessite des interactions langagières ; elle implique des relations interpersonnelles substantielles, avec une forte interdépendance. Il s'agit d'un travail d'équipe occasionnant un sentiment d'efficacité collective, relevant de ce que (Bandura, 2000) nomme l'indice agrégatif holistique.

3.3.5 Le travail de groupe

Avec le travail *de* groupe en revanche, le niveau d'interdépendance est faible entre les membres. Cependant il n'est pas inexistant : le travail individuel est coordonné en vue d'un objectif commun. L'indice utilisé pour mesurer le sentiment d'efficacité collective d'un tel groupe serait donc ce que (Bandura, 2000) nomme l'indice agrégatif personnel.

3.3.6 La notion de collectif apprenant

En particulier, (Lescouarch, 2021 : 8) souligne l'importance de la notion de « collectif apprenant » pour « trouver un équilibre efficace entre le travail collectif en classe entière et le travail individualisé des élèves » (Cnesco, 2017, p. 6). Il s'agit de construire un environnement personnalisé et étayant (Lescouarch, 2018), en appui sur un collectif qui déborde le cadre de la classe pour s'étendre à l'équipe enseignante. En mutualisant les pratiques et les ressources matérielles, en coanimant régulièrement des cours – ce qui est une forme de

« coformation » –, le collectif apprenant peut ainsi enrichir la différenciation pédagogique et favoriser l'inclusion de tous les élèves.

Ce qu'il faut retenir

Dans la présente étude, nous avons choisi d'employer le terme *coopération* lorsqu'il s'agit d'interactions entre apprenants ; tandis que nous avons réservé l'acception *collaboration* aux interactions entre les enseignants et le chercheur, *e.g.* selon le paradigme de la recherche collaborative en sciences de l'éducation et de la formation (SDEF). Du point de vue des apprenants, les pédagogies coopératives plongent leurs racines dans l'Antiquité et, aujourd'hui en France, elles occupent une place importante dans les textes officiels du ministère de l'Éducation nationale. Parmi les approches praxéologiques contemporaines en SDEF, une typologie a émergé, qui permet de distinguer entre plusieurs modalités ou approches coopératives – *e.g.* l'aide, l'entraide, le tutorat, le travail en groupe ou la notion de collectif apprenant –, qui impliquent toutes, à divers degrés, l'autodirection des apprentissages.

4 L'autodirection des apprentissages

Dans ce chapitre, nous abordons *l'autodirection des apprentissages* : dans le prolongement de ceux abordés dans les chapitres précédents, ce concept théorisé par (Carré, 2010) offre l'avantage d'articuler de façon opérationnelle deux concepts clés, en lien à la fois avec l'autonomie des apprenants et avec leur capacité à coopérer, à savoir : l'autorégulation des apprentissages (Efklides, 2011; Pintrich, 2000; Zimmerman, 2000) et l'auto-efficacité (Bandura, 2007).

Pour commencer nous allons examiner la motivation et la métacognition, en tant que concepts sous-jacents de l'autorégulation des apprentissages, puis nous examinerons la notion d'auto-efficacité.

4.1 L'autorégulation des apprentissages

Parmi les principaux modèles de l'autorégulation des apprentissages (« *self-regulated learning* » ou SRL en anglais), on peut citer celui de (Zimmerman, 2008). Selon cet auteur, l'autorégulation des apprentissages correspond aux processus par lesquels les élèves initient et

entretiennent les concepts, les émotions et les comportements propices à la réalisation de leurs objectifs d'apprentissage. Il propose une description descendante (« *top-down* ») des différentes stratégies cognitives et métacognitives mobilisées par l'apprenant pour exécuter une tâche. Citons également le modèle de (Pintrich, 2000) qui accorde à la dimension affective une place prépondérante dans les processus autorégulateurs. Pour finir, (Carré, 2010) souligne l'importance de « l'articulation de processus motivationnels et métacognitifs » (Carré *et al.*, 1997) dans l'autorégulation des apprentissages. Les sections qui suivent ont pour objectif d'explicitier cette articulation.

4.1.1 Le rôle de la motivation

4.1.1.1 Modèle intégratif de la motivation selon (Fenouillet, 2009)

Pour commencer il est intéressant de noter que, d'emblée, (Carré, 2010, p. 122) situe la notion d'autonomie comme étant à la source de l'Education nouvelle :

« Les thèmes de la liberté, de la responsabilité et de l'autonomie pour apprendre traversent l'histoire des idées pédagogiques, tant pour ce qui est de l'éducation initiale que pour ce qui concerne ce que l'on appelle aujourd'hui la formation tout au long de la vie. On peut même s'essayer à considérer l'éducation « nouvelle » comme aux sources de l'idée d'autoformation [...] les grands maîtres de la pédagogie (Jacotot, Pestalozzi, Cousinet, Freinet, etc.) ont toujours porté au fronton de leurs systèmes pédagogiques l'autonomie du sujet apprenant ».

Cette idée rejoint celle de (Cosnefroy & Fenouillet, 2009, pp. 130-131), selon qui, s'il est « difficile de donner des recettes ou des méthodes pour susciter l'intérêt ou pour faire en sorte que l'élève accorde de l'importance aux activités scolaires [...] il semble possible de prôner un principe sans lequel l'un et l'autre peuvent difficilement s'épanouir. Ce principe [étant] celui d'une recherche d'autonomie ».

Pour éclairer ensuite le lien entre motivation et autodirection des apprentissages, on peut s'appuyer sur la méta-analyse réalisée par (Fenouillet, 2009). Dans le prolongement de (Bandura, 2007), cet auteur met notamment en évidence l'articulation entre le sentiment d'efficacité personnelle et la motivation, « médiateur [de ce dernier] dans le processus qui permet à l'individu de produire les effets qu'il recherche ».(Fenouillet, 2009, p. 308). En outre, pour « comprendre toutes les composantes qui vont de l'origine jusqu'à l'aboutissement de

l'action » (Fenouillet, 2009, p. 333), une distinction utile est établie par des théoriciens de la motivation comme (Kuhl, 1987; Zimmerman, 2002) entre des aspects motivationnels au sens strict, *i.e.* qui interviennent en amont de la décision, versus des aspects qui interviennent en amont de la décision d'agir, et qui relèvent de ce que ces auteurs nomment, à l'instar de (Fenouillet, 2009), la « volition », *i.e.* une variation des buts ou des intentions, dans le cours même de l'action, qui résulte en une modification de la motivation.

En effet, selon (Cosnefroy, 2010), les apprenants doivent faire face à une problématique double : la mise au travail d'une part, l'entretien de la dynamique de travail d'autre part. Cet auteur distingue ainsi entre la motivation – par laquelle l'apprentissage est initié – et la volition, grâce à laquelle l'apprenant peut maintenir son effort. Cette articulation entre motivation et volition est explicitée par (Carré, 2010, p. 133) :

« La première dimension de l'autodirection recouvre alors les dimensions motivationnelles du “pourquoi” de l'engagement dans l'action, du désir et de la détermination des buts et des finalités, tandis que la seconde traduit les dimensions volitionnelles du “comment” de la conduite de l'action, des stratégies d'action, de la persistance du comportement et de sa régulation ».

4.1.1.2 Le rôle de la motivation dans les apprentissages scolaires

Pour commencer, (Cosnefroy & Fenouillet, 2009, pp. 129-130) rappellent, en s'appuyant sur (Ryan & Deci, 2002), la distinction entre motivation intrinsèque et motivation extrinsèque chez les apprenants : la motivation intrinsèque implique de s'engager librement dans une activité pour le plaisir qu'elle procure, tandis que la motivation extrinsèque implique la recherche d'une gratification extérieure à la tâche. (Cosnefroy & Fenouillet, 2009, p. 130) précisent cependant que cette distinction n'est pas nécessairement pertinente et opératoire en contexte scolaire, dans la mesure où « il existe des formes fortement autodéterminées de motivation extrinsèque ». A cette distinction entre motivation intrinsèque et extrinsèque, ces auteurs jugent donc préférable de substituer une distinction « entre motivation autonome qui implique choix personnel et motivation contrôlée qui suppose coercition et pression ». Cette question de l'autonomie des apprenants joue donc selon eux un rôle central pour la motivation des élèves.

4.1.1.3 Le continuum motivationnel de l'autodétermination

En s'appuyant sur (Ryan & Deci, 2000), (Carré, 2010, p. 142) affine cette dichotomie entre motivations extrinsèque et intrinsèque.

Tout d'abord, la motivation intrinsèque est considérée comme le principe essentiel de ce qu'il nomme *l'autodétermination* dans la mesure où elle implique que l'apprenant entreprenne une action ou adopte un comportement de sa propre initiative, sans qu'une stimulation ou une contrainte extérieure soit nécessaire. Une action ou un comportement intrinsèquement motivés se traduisent par une forme de gratuité : ils sont gratifiants par eux-mêmes aux yeux de l'apprenant, *e.g.* parce qu'ils satisfont sa curiosité intellectuelle ou son goût pour l'altruisme. En contexte scolaire, il peut exister une tension entre cette motivation intrinsèque et la motivation extrinsèque structurellement induite par la nécessité d'obtenir de bonnes notes ou un diplôme par exemple.

Concernant la motivation extrinsèque, (Carré, 2010; Ryan & Deci, 2000) distinguent cependant plusieurs formes, dont certaines sont autodéterminées. D'une part la *motivation extrinsèque contrôlée* ou *par régulation externe* implique que l'apprenant agisse sous l'effet (ou sous la menace) de sanctions externes – positives ou négatives. A cet égard, la note joue souvent un rôle de motivation extrinsèque contrôlée en contexte scolaire. Selon cette typologie, les auteurs précités évoquent ensuite la *motivation extrinsèque introjectée*, qui correspond à une motivation extrinsèque où la régulation externe est davantage intériorisée par le sujet. Certains élèves dont le comportement est considéré comme « scolaire » par leurs enseignants, peuvent être motivés de la sorte. Une première forme de motivation autodéterminée selon (Ryan & Deci, 2000) est ce qu'ils nomment la *motivation extrinsèque identifiée* : l'apprenant s'émancipe du contrôle extérieur pour agir d'après son propre jugement. Cette émancipation procède par identification dans la mesure où les valeurs ou les intérêts d'autrui sont considérés comme compatibles avec ceux du sujet. Enfin, (Carré, 2010) évoque la *motivation extrinsèque intégrée*, qui constitue selon (Ryan & Deci, 2000) la forme de motivation extrinsèque la plus autodéterminée, dans la mesure où l'apprenant y agit en étant pleinement conscient de son autonomie décisionnelle, et en accord total avec ses valeurs profondes. En contexte scolaire, cette motivation extrinsèque intégrée semble difficile à favoriser ; à moins d'envisager par exemple des pédagogies coopératives laissant à l'apprenant une large part d'autonomie.

A partir de ce qui précède, (Carré, 2010; Ryan & Deci, 2000) mettent en évidence un continuum motivationnel, allant de l'absence de motivation (*i.e. amotivation*) à la motivation intrinsèque, en passant par les motivations extrinsèques contrôlée, introjectée, identifiée et

intégrée – classées par ordre croissant d'autodétermination. La Figure 4 résume ce continuum motivationnel de l'autodétermination selon (Ryan & Deci, 2000), rapporté par (Carré, 2010).

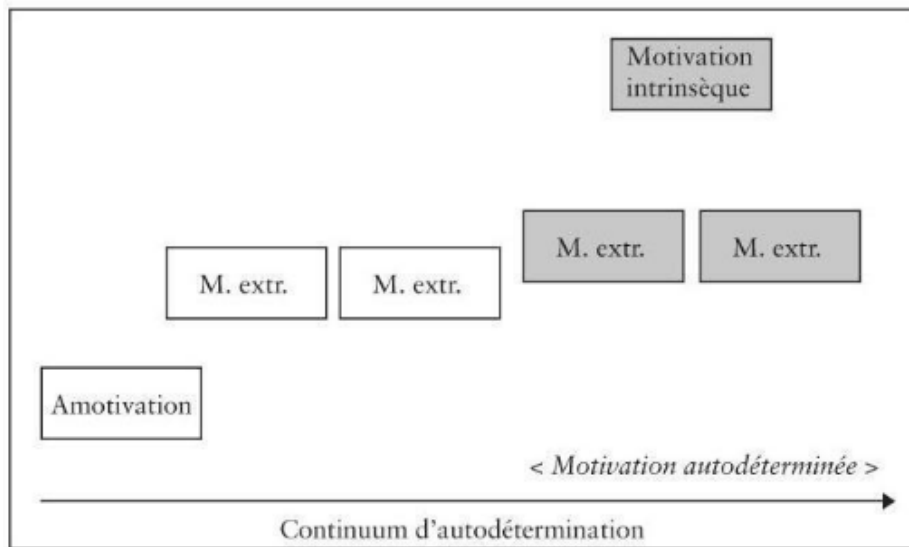


Figure 4 – Continuum de l'autodétermination selon (Ryan & Deci, 2000) rapporté par (Carré, 2010)

4.1.1.4 Motivation et auto-efficacité

Il semble néanmoins qu'un autre facteur doive être pris en compte pour initier puis maintenir l'autonomie et la motivation des apprenants, à savoir le *sentiment d'efficacité personnelle* (Bandura, 2007). Relativement au concept intégratif de l'autodirection des apprentissages selon (Carré, 2010), le lien entre motivation et auto-efficacité peut d'ailleurs être mieux compris grâce à (Cosnefroy & Fenouillet, 2009, p. 132), qui évoquent l'importance de « la perception de compétence » pour la motivation et l'autonomie. Cette notion, qui est à rapprocher du sentiment d'efficacité personnelle, est distincte de « la notion de compétence qui est liée aux différents savoirs que possède l'individu » : en effet elle concerne plutôt « l'appréciation, [le] jugement que porte l'élève sur ce qu'il pense avoir comme compétence et ce, indépendamment d'une évaluation (plus ou moins) objective qui pourrait être produite par ailleurs » (*ibid.*).

4.1.2 Le rôle de la métacognition

Pour comprendre le rôle de la métacognition dans les apprentissages autodirigés, on pourra notamment s'appuyer sur les travaux de (Grangeat, 1999) et sur la méta-analyse réalisée par (Berger & Büchel, 2012).

4.1.2.1 Pour définir la métacognition

Pour tenter de définir la métacognition, on peut commencer par s'appuyer sur une définition proposée par (Flavell, 1976), dont la définition fait référence, selon (Dortier, 2014, p. 25) ; en effet, ce psychologue est « l'un des pionniers des recherches sur le sujet ». La métacognition serait ainsi « la connaissance que l'on a de ses propres processus cognitifs ». Cette définition rejoint la formulation employée par (Grangeat, 1999, p. 116), citant (Romainville *et al.*, 1995), et selon qui la métacognition serait une « opération mentale portant sur d'autres opérations mentales ». De même, dans le dictionnaire de l'*American Psychological Association* (VandenBos, 2007, p. 645), on peut lire que la métacognition correspond à la « conscience qu'à un sujet de ses propres processus cognitifs, ce qui implique souvent un effort conscient pour les contrôler ». On voit ici apparaître deux dimensions distinctes mais complémentaires. De fait, si (Dortier, 2014, p. 25) signale que le mot *métacognition* peut sembler, de prime abord, « un terme complexe qui désigne une idée simple : apprendre à apprendre » ; il n'en demeure pas moins que, pour (Gombert, 1990, p. 27), cité par (Romainville *et al.*, 1995, p. 45), la métacognition est effectivement un concept *composite* qui regroupe, d'une part, « les connaissances introspectives conscientes qu'un individu particulier a de ses propres états et processus cognitifs » et d'autre part, « les capacités que cet individu a de délibérément contrôler et planifier ses propres processus cognitifs en vue de la réalisation d'un but ou d'un objectif déterminé ».

Pour (Grangeat, 1999), cette première facette de la métacognition a trait aux connaissances de l'apprenant concernant les stratégies cognitives efficaces pour réaliser efficacement une tâche : c'est ce que l'auteur appelle les *métaconnaissances*. L'autre dimension de la métacognition, que (Grangeat, 1999, p. 115) appelle *régulation métacognitive*, concerne la capacité de l'apprenant à anticiper des procédures, à les évaluer et à les adapter pour mener à bien ses apprentissages.

4.1.2.2 Métacognition et différenciation pédagogique selon (Grangeat, 1999)

Selon (Grangeat, 1999), des activités métacognitives se traduiraient, dans la classe, par des dispositifs mettant en œuvre la différenciation pédagogique et caractérisées par une « conduite réfléchie des démarches d'apprentissage » requérant trois conditions : premièrement que l'élève connaisse les objectifs de l'enseignement (« enseignement explicite »), qu'il dispose d'un degré d'autonomie pour choisir entre plusieurs démarches d'apprentissage possibles et qu'il puisse connaître les effets de ses choix, deuxièmement qu'il

parvienne à mobiliser différents « modes de formulation ou de schématisation [des informations] », et enfin qu'il puisse « comparer ses propres stratégies cognitives à celles de ses pairs ou s'approprier celles d'un médiateur » (Grangeat, 1999, p. 126).

4.1.2.3 Articulation entre métacognition et motivation

Dans leur méta-analyse, (Berger & Büchel, 2012, p. 96) s'appuient notamment sur (Efklides, 2006; Pintrich *et al.*, 2000) pour distinguer trois composantes de la métacognition : « *connaissances* métacognitives, *expériences* métacognitives et *stratégies* métacognitives ».

Pour commencer, les connaissances métacognitives correspondent aux « métaconnaissances » que nous avons déjà rencontrées plus haut avec (Grangeat, 1999). (Berger & Büchel, 2012) précisent toutefois que ces métaconnaissances sont stockées dans la mémoire à long-terme de l'apprenant, et qu'elles peuvent être de trois types (Paris *et al.*, 1983, p. 303) : déclaratives (savoirs), procédurales (savoir-faire) ou conditionnelles (savoir quand et pourquoi recourir à une stratégie). Cependant, selon (Büchel, 1991) ou en s'appuyant sur la définition que donne (VandenBos, 2007, p. 1036) du mot « stratégie »¹⁴, on peut remettre en question la pertinence de cette troisième catégorie de métaconnaissances, dans la mesure où on peut considérer que toute stratégie est intrinsèquement procédurale et inclut ainsi une connaissance des circonstances propices à son application.

Les expériences métacognitives constituent quant à elles « des sentiments et jugements subjectifs » en lien avec le(s) processus cognitif(s) en cours lors de la réalisation d'une activité d'apprentissage. Ces expériences surviennent quand l'élève « prend conscience de ses processus de pensée actuels » et leur rôle principal est de « guider les processus de pensée d'un point de vue subjectif et du point de vue des connaissances spécifiques à la tâche, ceci afin d'influencer les stratégies métacognitives » (Efklides, 2001, cité par Berger & Büchel, 2012, p. 96). Les jugements subjectifs formulés par l'apprenant dans le cadre de ses expériences métacognitives concernent la difficulté estimée de la tâche ou la confiance relative dans l'exactitude d'une solution que l'élève a trouvée pour celle-ci. Ces expériences peuvent aussi prendre « la forme de sentiments tels que le sentiment de familiarité avec la tâche » (*ibid.*).

Troisièmement, les stratégies métacognitives correspondent aux choix opérés par l'apprenant, en se fondant consciemment ou inconsciemment sur ses expériences métacognitives, d'une part pour anticiper et planifier l'activité en amont de la résolution d'une tâche, et d'autre part pour vérifier et adapter les processus cognitifs en aval de la réalisation

¹⁴ « Programme d'action conçu pour réaliser un objectif ou accomplir une tâche » (traduction libre)

d'une tâche. « Les expériences métacognitives et les stratégies métacognitives sont [donc] fortement interdépendantes » (Berger & Büchel, 2012).

Plus largement, c'est l'interdépendance entre métacognition et motivation qui est soulignée par des auteurs comme (Berger & Büchel, 2012; Zimmerman, 1995) pour favoriser les apprentissages des élèves. Notamment parce que « des croyances motivationnelles défavorables sont liées à un moindre exercice de stratégies métacognitives ainsi qu'à des jugements métacognitifs moins adaptés » (Berger & Büchel, 2012, p. 117). Ainsi, pour mieux analyser les interactions entre motivation et métacognition, (Zimmerman, 1995) propose explicitement d'opérationnaliser cette notion de « croyance motivationnelle » au moyen du sentiment d'efficacité personnelle de (Bandura, 2007). De même, (Fenouillet, 2009, pp. 307-308) affirme que, bien que la théorie du sentiment d'efficacité personnelle de Bandura « ne se réclame pas spécialement du champ de la motivation », elle s'avère pourtant « utile et pertinente pour la définir ou l'expliquer ».

Les expériences métacognitives jouent par ailleurs « un rôle médiateur entre croyances motivationnelles et engagement cognitif » : c'est pourquoi (Berger & Büchel, 2012, p. 117) suggèrent que l'enseignement explicite – qui consiste à « présenter la matière de façon fractionnée, à vérifier la compréhension, et à s'assurer d'une participation active et fructueuse de tous les élèves » (Gauthier *et al.*, 2007, p. 2) – pourrait constituer un modèle pertinent « pour développer des croyances motivationnelles facilitant l'apprentissage, l'utilisation et le transfert des stratégies d'apprentissage cognitives et métacognitives » (Berger & Büchel, 2012, p. 117).

De même, ces auteurs mettent en avant les travaux de (Schunk, 1982), qui a expérimenté ce modèle d'enseignement explicite en mettant en lumière le rôle bénéfique du feedback pour les apprentissages des élèves : un dispositif qui s'est révélé fructueux pour le « développement du sentiment d'efficacité personnelle chez des élèves rencontrant des difficultés d'apprentissage » (Berger & Büchel, 2012, p. 119). Par ailleurs, selon ces auteurs, le sentiment d'efficacité personnelle peut également bénéficier du *modelage*, *i.e.* lorsque l'enseignant lui-même se donne en exemple dans la réalisation d'une tâche, explicitant ses stratégies métacognitives pour que les élèves puissent se les approprier. Or les travaux de (Schunk & Ertmer, 2000) ont montré que cette pratique peut être d'autant plus profitable pour les apprenants – notamment ceux qui sont le plus en difficulté – s'ils modèlent leurs stratégies métacognitives d'après celles de leurs pairs qui sont davantage en réussite.

En définitive, ces auteurs formulent une préconisation méthodologique pour la recherche dans ce champ, en faveur d'études longitudinales mesurant à plusieurs reprises sur une certaine période les croyances motivationnelles (*e.g.* le sentiment d'efficacité personnelle)

et les aspects métacognitifs, « afin d’appréhender la causalité réciproque telle qu’évoquée dans les modèles théoriques ».

Après avoir ainsi tenté d’éclairer la dynamique de l’action de l’apprenant, au moyen de cette relation dialectique « entre facteurs “motivationnels” et “métacognitifs” de l’autodirection [des apprentissages] » (Carré, 2010, p. 133), nous allons maintenant pouvoir nous intéresser à la notion d’*auto-efficacité*.

4.2 L’auto-efficacité, « coefficient » de l’autodirection des apprentissages

4.2.1 Le sentiment d’efficacité personnelle

Le concept d’auto-efficacité est emprunté au champ de la psychologie cognitive et dérivé de la théorie de l’agentivité humaine selon Albert Bandura, qui est considéré comme l’un des plus importants chercheurs en psychologie au XXI^e siècle. Ses contributions séminales dans le champ de la psychologie cognitive notamment, le placent selon certains au même rang que des chercheurs influents comme Piaget ou Vygotsky, avec les théories desquels, selon Tudge & Winterhoff (1993), sa propre théorie de l’agentivité humaine entretient une proximité théorique significative.

Le sentiment d’efficacité personnelle (Bandura, 2007) d’un individu peut être défini comme sa croyance en sa capacité à accomplir une tâche avec succès ou à influencer efficacement sur son environnement. Ce déterminant majeur de l’action humaine (Bandura, 2009; Bandura *et al.*, 1996) – particulièrement pour le XXI^e siècle (Carré, 2004) – ne concerne donc pas tant le nombre de compétences qu’un individu maîtrise, mais « ce qu’il croit pouvoir en faire dans des situations variées » : le SEP ou « auto-efficacité perçue » constitue donc plutôt « une capacité productrice » pouvant faciliter une mobilisation efficace « des sous-compétences cognitives, sociales, émotionnelles et comportementales (...) pour servir de nombreux buts » (Lecomte, 2004, p. 60).

4.2.2 Le sentiment d’efficacité collective

4.2.2.1 Définition

Le sentiment d’efficacité collective est le pendant groupal du sentiment d’efficacité personnelle. Il peut être opérationnalisé avec profit en sciences de l’éducation et de la

formation, dans une perspective psychopédagogique – ou « psycho-andragogique » si l'on se réfère à la formation des adultes –, notamment pour étudier des situations d'apprentissages collaboratifs (Bellet *et al.*, 2019).

4.2.2.2 Indice agrégatif personnel vs. indice agrégatif holistique

Selon (Bandura, 2000, 2006), il existe deux approches bien distinctes pour mesurer le sentiment d'efficacité collective, selon la situation concernée. D'une part, l'indice agrégatif personnel (IAP) consiste en l'agrégation des mesures des perceptions des membres du groupe concernant leur capacité personnelle à accomplir les fonctions spécifiques qui leur incombent au sein de celui-ci. D'autre part, l'indice agrégatif holistique (IAH) consiste en l'agrégation des mesures des perceptions de chaque individu concernant la capacité de leur groupe à opérer comme un ensemble cohérent : il s'agit d'une approche holistique, englobant les aspects de coordination et d'interaction qui opèrent au sein des groupes.

En d'autres termes, si l'activité collective à mesurer s'apparente au jeu collectif d'une équipe de football, où l'interdépendance est élevée, l'indice utilisé sera l'IAH. En revanche, si l'activité collective à mesurer s'apparente aux performances individuelles d'une équipe de gymnastique aux jeux olympiques, où l'interdépendance est basse, l'indice utilisé sera l'IAP.

4.2.2.3 Comment mesurer le SEC ?

Pour mesurer le SEC, plusieurs pistes peuvent être envisagées, *e.g.* des emprunts à (Cosnefroy & Jézégou, 2013) ou (Wang & Lin, 2007).

Dans le cas de Cosnefroy & Jézégou (2013), l'échelle utilisée est en fait une sous-échelle comportant 3 items, issue d'une échelle composite plus large (20 items), création ad hoc des auteurs pour mesurer l'autorégulation collective et individuelle des apprentissages, dans une situation de pédagogie par projet, avec des étudiants ingénieurs effectuant un travail de groupe :

- Je pense que notre groupe de travail est performant.
- Je pense que notre groupe obtiendra une bonne évaluation à la fin de cette phase.
- Par rapport à la réussite finale du projet, le groupe se sent confiant.

Dans le cas de (Wang & Lin, 2007), il s'agit d'une transposition réalisée à partir des items d'une échelle mesurant le sentiment d'efficacité personnelle, de manière à ce qu'ils s'appliquent au travail de groupe. Ici aussi, il s'agit donc d'une approche holistique – *i.e.* mesurée au moyen de l'IAH :

« [Notre] échelle mesurant le sentiment d'efficacité collective a été transposée à partir d'une échelle développée par l'auteur originel. Ainsi que le suggérait (Bandura, 2007), l'efficacité collective se rapporte à la capacité du groupe, considéré comme un tout, à performer. De ce fait, les items se rapportant à l'auto-efficacité ont été transposés du niveau individuel au niveau groupal. Par exemple, l'item "J'ai confiance en ma capacité à réaliser de l'excellent travail avec les devoirs et les examens dans ce cours" a été transposé en "J'ai confiance en la capacité de mon groupe à réaliser de l'excellent travail avec les devoirs et les examens dans ce cours". » (Wang & Lin, 2007, p. 2260, traduction libre).

Pour mesurer le sentiment d'efficacité collective selon l'indice agrégatif personnel (IAP), il est en revanche difficile de trouver des échelles dans la littérature.

4.2.3 Auto-efficacité et apprentissage tout au long de la vie

Bandura met par ailleurs l'accent sur la nécessité d'une école adaptée aux nouvelles exigences de la formation tout au long de la vie :

« Dans une société globalisée du savoir, qui change rapidement, l'espoir et l'avenir des gens réside dans leur capacité à se développer et à se renouveler par eux-mêmes continuellement. Les systèmes éducatifs doivent dorénavant mettre l'accent, non plus sur la seule transmission du savoir, mais sur le fait d'apprendre aux élèves à s'éduquer eux-mêmes tout au long de leur vie. Ces derniers doivent apprendre à s'adapter et devenir des apprenants qui s'autodirigent efficacement. L'éducation à l'autodirection est désormais vitale pour une société productive et innovante » (Bandura, 2002, p. 4, traduction libre).

L'accent que (Bandura, 2002) met ici sur l'importance du développement de l'autonomie des apprenants fait écho aux travaux de (Mailles Viard Metz, 2015b; Mailles Viard Metz *et al.*, 2014), notamment pour le développement d'outils numériques favorisant l'autonomie, à destination des formateurs et des étudiants (Mailles Viard Metz, 2015a). En particulier, un outil comme le e-portfolio offre des perspectives intéressantes pour la construction du projet d'orientation des étudiants (Loisy *et al.*, 2011).

4.2.4 L'auto-efficacité selon l'enquête PISA 2018

Nous avons déjà rencontré ce concept dans l'enquête PISA 2018. Voici comment les auteurs de cette étude se réfèrent à la littérature en psychologie de l'éducation pour justifier leur recours au sentiment d'efficacité personnelle (OCDE, 2019b, p. 188, traduction libre) :

« Le regard que les apprenants portent sur leurs capacités, et leur peur de l'échec, peut influencer leurs émotions, leur motivation et leur comportement. Selon la théorie de l'agentivité humaine, les apprenants sont davantage susceptibles de se fixer des objectifs ambitieux, de fournir plus d'efforts et persévérer plus longtemps lorsqu'ils croient qu'ils vont réussir. Et réciproquement (...). Il est probable que les apprenants dont l'auto-efficacité perçue est plus basse ne réaliseront pas pleinement leur potentiel, et compromettront leur scolarité, ainsi que leurs objectifs en termes de carrière ».

Or les résultats de cette enquête PISA montrent que « les élèves en France expriment un plus faible sentiment d'auto-efficacité [...] que la moyenne observée dans les pays de l'OCDE ».

4.2.5 Le sentiment d'efficacité personnelle des enfants (SEPE)

Dans une vaste étude transnationale impliquant 3 pays, (Pastorelli *et al.*, 2001) ont quant à eux enquêté auprès de 1180 écoliers d'un âge moyen de 12,5 ans, afin d'étudier les déterminants du sentiment d'efficacité personnelle des enfants/pré-adolescents. Pour ce faire, ils ont utilisé le questionnaire SEPE, une échelle psychométrique multidimensionnelle construite d'après les recommandations de (Bandura, 2006). Cette échelle comporte 37 items et mesure le sentiment d'efficacité personnelle des enfants, pré-adolescents ou adolescents au moyen d'une échelle de Likert à cinq points.

(Bandura, 2007) s'appuie sur les travaux de (Pajares *et al.*, 2001) pour rappeler que dans le contexte des échelles mesurant le sentiment d'efficacité personnelle, des échelles de Likert comportant un plus grand nombre de points permettent aux répondants de mieux discriminer leurs réponses, ce qui renforce *in fine* la prédictivité. Il recommande ainsi d'utiliser une échelle de Likert à 10 ou 100 points.

Avec un public jeune toutefois, une échelle de Likert comportant un trop grand nombre de points peut être source de confusion et nuire aux résultats. Dans leur étude de référence,

(Pastorelli *et al.*, 2001) ont donc opté pour un format à 5 points. De même, (Masson & Fenouillet, 2013) ont utilisé une échelle de Likert à six points pour leur étude portant sur le sentiment d'efficacité personnelle d'élèves de cycle 3, auprès d'un échantillon comprenant 301 élèves (148 filles et 153 garçons) de CM1-CM2 de 17 classes différentes (âge moyen 11,39 ans).

L'échelle SEPE est par ailleurs multidimensionnelle dans la mesure où elle comporte 7 dimensions (Pastorelli *et al.*, 2001, p. 88), pouvant être considérées comme des *sous-échelles* :

1. L'auto-efficacité perçue pour la réussite scolaire, qui mesure les représentations des élèves dans leur capacité à maîtriser des compétences dans différentes disciplines scolaires [SEP(scol)] ;
2. L'auto-efficacité perçue pour l'autorégulation des apprentissages, qui évalue les représentations des élèves en leur capacité à structurer leur environnement pour le rendre propice aux apprentissages, ainsi qu'à planifier et organiser des activités académiques [SEP(srl)] ;
3. L'auto-efficacité perçue pour les compétences liées aux loisirs et aux activités périscolaires, qui mesure les représentations des élèves en leur capacité à s'adonner à des activités récréatives et à prendre part à des travaux d'élèves en groupe ;
4. L'auto-efficacité perçue pour l'autorégulation, qui mesure les représentations des élèves en leur capacité à résister à la pression exercée par les pairs relativement à des activités à haut risque impliquant l'alcool, la drogue ou des conduites transgressives¹⁵ ;
5. L'auto-efficacité perçue pour les apprentissages coopératifs, qui mesure les représentations des élèves en leur capacité à initier et cultiver des relations sociales et à gérer les conflits interpersonnels [SEP(coop)] ;
6. L'auto-efficacité perçue pour l'affirmation de soi, qui mesure les représentations des élèves en leur capacité à exprimer leurs opinions, à ne pas se laisser faire en cas de mauvais traitements et à refuser les demandes déraisonnables ;

¹⁵ Bien que cette dimension ne soit pas pertinente avec un public d'enfants, elle constitue un élément important de l'échelle SEPE lorsqu'elle est utilisée avec des pré-adolescents ou à fortiori avec des adolescents. Pour rappel, l'âge moyen des enquêtés dans l'étude de (Pastorelli *et al.*, 2001) est 12,5 ans, un âge auquel ces problématiques peuvent constituer, sinon une réalité, du moins une virtualité (Leslie, 2008; Oubrayrie-Roussel & Safont-Mottay, 2001).

7. L'auto-efficacité perçue pour répondre aux attentes d'autrui, qui mesure les représentations des élèves en leur capacité à satisfaire ce que leurs parents, leurs enseignants et leurs pairs attendent d'eux, et à satisfaire aux attentes qu'ils ont vis-à-vis d'eux-mêmes.

Dans l'étude de (Pastorelli *et al.*, 2001), les 37 items originellement formulés en langue anglaise par (Bandura, 2006, pp. 326-327) font l'objet d'un processus de traduction et de rétro-traduction en double aveugle par deux locuteurs bilingues indépendants, avant une harmonisation permettant d'aboutir à une version de travail consensuelle (Pastorelli *et al.*, 2001, p. 89).

(Bandura, 2005, p. 315) recommande en outre de prétester l'échelle SEPE pour en vérifier la validité interne au moyen de l'alpha de Cronbach et éliminer ou reformuler les items ambigus.

4.3 Autodirection et apprentissages de préadolescents

Nous avons vu que l'autodirection des apprentissages articule autorégulation et auto-efficacité ; on peut toutefois se demander dans quelle mesure ce concept élaboré par (Carré, 2010) dans le champ de l'andragogie et dans le contexte de la formation pour adultes, pourrait s'appliquer au champ de la pédagogie, en contexte scolaire – *e.g.* avec un public de collégiens (préadolescents).

D'emblée, citant (Schwartz, 1973) et (Dumazedier, 1985), Carré insiste sur le fait que l'autodirection des apprentissages constitue pour « le système éducatif » un enjeu important « dès la jeunesse » (Poisson *et al.*, 2015, p. 119). L'auteur précise que ce sont les travaux de Cosnefroy, Fenouillet, Deci, Ryan, Zimmerman et Bandura notamment, qui « [l'autorisent] à en présenter une synthèse conceptuelle et opératoire » : *i.e.* le concept intégratif d'*autodirection des apprentissages*. Or Cosnefroy et Fenouillet, deux auteurs dont les méta-analyses (Cosnefroy, 2011; Fenouillet, 2009) ont effectivement constitué pour Carré (2010) un point d'appui important, ont travaillé leurs construits théoriques à la fois avec des publics adultes et des publics d'enfants, préadolescents ou adolescents en milieu scolaire (Cosnefroy & Fenouillet, 2009) : leurs théories, respectivement, de la motivation et de l'autorégulation des apprentissages peuvent se montrer pertinentes dans les deux contextes.

En outre, les autres auteurs mentionnés par Carré (2010), auteurs anglo-saxons de référence dans leurs champs respectifs, et sur lesquels il s'appuie pour articuler le concept intégratif d'autodirection des apprentissages, ont élaboré des théories qui s'appliquent à

l'apprenant en général et n'excluent pas explicitement l'apprenant préadolescent de leur champ d'investigation. C'est notamment le cas de (Zimmerman, 1995, 2000, 2002, 2008) pour sa théorie de l'autorégulation des apprentissages et de (Deci & Ryan, 2008; Ryan & Deci, 2000, 2002) pour leur théorie de la motivation.

Quant à l'auto-efficacité, (Bandura, 2007, pp. 326-363) consacre à « l'efficacité personnelle cognitive des élèves » une section entière de son ouvrage de référence. Il y aborde notamment le rôle de l'efficacité perçue dans le développement cognitif autorégulé des élèves, en s'appuyant sur des études empiriques menées de l'école primaire au lycée, avec des enfants, des préadolescents et des adolescents. Le célèbre psychologue affirme par ailleurs que « les enfants et les adultes accomplissent bien plus de choses grâce à l'influence autorégulatrice que sans elle » (Bandura, 2007). Cette expression englobante montrant encore une fois que les concepts liés à l'autodirection des apprentissages peuvent s'appliquer à la fois aux adultes et aux enfants (Bandura, 1986). L'auteur a d'ailleurs consacré un article à la question de l'auto-efficacité perçue chez les adolescents (Bandura, 2006).

Il est intéressant de noter que Carré (2010) lui-même n'exclut pas explicitement les enfants, préadolescents et adolescents du champ d'application de l'autodirection des apprentissages. Au contraire, il met en avant les travaux de recherche de certains de ses collaborateurs qui ont investi ce champ dès l'école primaire. C'est notamment le cas de (Bui, 2007), dont il a dirigé la thèse de doctorat, et qui a opérationnalisé l'autodirection des apprentissages avec succès dans le cadre d'une étude empirique menée auprès d'enfants en milieu scolaire (Poisson *et al.*, 2015, p. 120). Carré (2010) insiste en outre sur le fait que l'on peut « considérer l'éducation "nouvelle" comme aux sources de l'idée d'autoformation » (Poisson *et al.*, 2015, p. 122 ; Carré, 2003) ; or l'Education nouvelle a largement investi le contexte scolaire dès ses origines, dans les années 1900, avec des publics d'enfants, de préadolescents et d'adolescents.

En somme, bien que l'autodirection des apprentissages soit le plus souvent associée au champ de la formation pour adultes, il n'en demeure pas moins que les théories qu'il articule sont également considérées par leurs auteurs comme pertinentes et applicables pour l'enseignement/apprentissage avec des enfants, des préadolescents ou des adolescents.

Ce qu'il faut retenir

L'autodirection des apprentissages est un concept intégratif articulant autorégulation des apprentissages, métacognition et auto-efficacité : trois concepts qui entretiennent des relations étroites et constituent des déterminants de la différenciation pédagogique. La motivation peut être appréhendée en tant que continuum, selon qu'elle est plus ou moins extrinsèque ou intrinsèque. La métacognition entretient avec la motivation une relation dialectique, de telle sorte que l'une ne peut pas être envisagée sans l'autre. Or l'auto-efficacité s'avère déterminante pour les mécanismes de la dyade motivation/métacognition ; cependant les élèves français semblent manifester une auto-efficacité moindre qu'ailleurs dans l'OCDE. Enfin, bien que l'autodirection des apprentissages ait été conceptualisée dans le champ de l'andragogie, ses concepts sous-jacents s'avèrent applicables au champ de la pédagogie, et permettent de mieux appréhender les enjeux des pédagogies coopératives en général, notamment dans la perspective de l'évaluation, qui fait l'objet du prochain chapitre.

5 Vers une évaluation coopérative

Pour rappel, après avoir montré que les pédagogies coopératives constituaient un moyen privilégié pour opérationnaliser la différenciation pédagogique – approche la plus souvent mise en avant pour répondre au défi des hétérogénéités et de l'accroissement des inégalités socioscolaires au collège –, nous avons vu que l'autodirection des apprentissages permettait de mieux comprendre les ressorts psychopédagogiques de la coopération. A travers le présent chapitre, nous essaierons de voir dans quelle mesure une réflexion sur les pratiques évaluatives en classe peut constituer un levier efficace pour mettre les approches précitées en œuvre.

5.1 Les quatre finalités évaluatives

En situation d'enseignement/apprentissage, l'évaluation est à la fois un concept clé et un sujet délicat, voire tabou (Chartier, 2013; De Landsheere, 1971). C'est également un sujet complexe, qui implique de prime abord qu'on établisse une distinction selon la finalité de l'évaluation (Cardinet, 1988; Perrenoud, 2001), qui peut être diagnostique (prise d'information), formative (appropriation ou entraînement), sommative (*e.g.* contrôle de fin de séquence) ou certificative (*e.g.* DNB ou Baccalauréat).

5.1.1 Le paradigme intuitif pragmatique

Lorsqu'une séquence d'enseignement est planifiée, avec les moments d'évaluation afférents, la prise en compte de ces quatre finalités peut s'avérer très importante. Or l'acte évaluatif est bien souvent déterminé par l'intuition et les exigences du contexte, avec une prédominance de la finalité *sommative*. Au préalable, l'enseignant n'a pas nécessairement précisé, ni ses objectifs, ni ses critères de réussite, ni la manière dont il utilisera les résultats de l'évaluation. De plus, là où les contraintes de temps liées à la prescription institutionnelle peuvent être ressenties comme fortes (« Il faut boucler le programme »), la préparation des situations d'évaluation peut être jugée secondaire par l'enseignant. Dans un tel contexte, l'attribution d'une note peut constituer un moyen pragmatique, apparemment objectif, d'informer les parents sur les performances de leurs enfants. Cependant ce système de notes peut dispenser les enseignants de « justifier les contenus précis des épreuves » (De Ketele, 1993). En outre, si en apparence la finalité d'un tel système est certificative, selon (Perrenoud, 1998), cité par (Flores *et al.*, 2015), de telles pratiques évaluatives ont avant tout une fonction, non nécessairement conscientisée, de gestion sociale, *e.g.* : gestion de la classe en vertu du pouvoir de motivation extrinsèque de la note, gestion institutionnelle eu égard au jugement que les collègues, les personnels de direction ou d'inspection pourraient porter sur les enseignants en fonction des notes qu'ils donnent à leurs élèves, ou gestion des relations parents-professeurs. En bref, les pratiques évaluatives les plus courantes, que (De Ketele, 1993) désigne sous l'appellation de *paradigme intuitif pragmatique*, seraient déterminées par une fonction non avouée et plus ou moins inconsciente de gestion sociale, tandis que les finalités diagnostique, formative, sommative et même certificative n'occuperaient que des rangs annexes parmi les préoccupations des enseignants.

5.1.2 De l'ère du test à l'ère de l'évaluation

Cependant depuis la fin du XXe siècle, une manière alternative d'envisager l'évaluation a progressivement émergé (Dochy, 2001), au point qu'à l'échelle du globe les systèmes éducatifs seraient peu à peu passés d'une *ère du test* à une *ère de l'évaluation* (Segers *et al.*, 2003b). L'ère du test est plutôt caractérisée par une séparation étanche entre moments réservés à l'instruction et phases de contrôle des connaissances, par une relative passivité des apprenants lors de ces dernières et par une tendance à ne mesurer les performances qu'au moyen d'un score total unique. À l'ère de l'évaluation, l'accent serait davantage mis sur l'articulation entre évaluation et instruction. Selon ce paradigme à tendance constructiviste, l'apprenant est vu

comme une personne active, qui prend ses responsabilités, réfléchit, collabore et entretient un dialogue continu avec l'enseignant (Dochy *et al.*, 1999). De ce fait, c'est aussi la finalité de l'évaluation qui est amenée à évoluer : au-delà de la fonction sommative (*e.g.* les notes dans le bulletin trimestriel) ou certificative (*e.g.* la délivrance du diplôme de fin d'étude), l'évaluation peut devenir un outil de suivi des progrès des élèves, afin de les orienter si nécessaire vers des activités de remédiation. Ainsi, dès la fin des années 1990, on trouve dans la littérature un nombre grandissant de publications qui, à l'instar de celles de (McIntosh, 1997; Pryor & Torrance, 1997), mettent l'accent sur les apports de la finalité formative de l'évaluation : l'évaluation comme outil pour apprendre, et non plus simplement comme une activité disjointe n'intervenant qu'une fois les apprentissages terminés. Cette approche rejoint celle de (Topping, 1998, p. 249), selon qui « l'évaluation formative a pour but d'améliorer les apprentissages pendant qu'ils ont lieu, afin de maximiser la réussite, plutôt que de simplement déterminer la réussite ou l'échec après-coup ». À l'usage, ainsi que l'ont signalé (Laveault & Allal, 2016), des synergies peuvent également se manifester entre les différentes finalités évaluatives. Certains enseignants peuvent par exemple instrumentaliser dans un but formatif des données initialement collectées dans un but sommatif (Black *et al.*, 2005). De même, une étude de (Brookhart, 2001) a montré comment des lycéens tendent à utiliser des indicateurs de type sommatif (*i.e.* des notes) pour autoréguler leurs apprentissages. En complément de la question de la finalité évaluative, un autre aspect important à prendre en compte est la question des *acteurs* du processus évaluatif, et de la *relation* qu'ils entretiennent. On aboutit ainsi à une typologie distinguant quatre modalités évaluatives, que nous allons maintenant expliciter.

5.2 Les trois principales modalités évaluatives

Partant du principe que, dans une salle de classe, il y a essentiellement deux types d'agents : l'apprenant et l'enseignant, on peut envisager trois modalités d'évaluation.

5.2.1 L'hétéro-évaluation

La modalité évaluative la plus courante – et peut-être parfois la seule qui soit pratiquée ? – consiste en l'évaluation d'un apprenant par un enseignant. Dans cette situation, l'enseignant, en tant que professionnel, est considéré comme un expert, compétent pour porter un jugement valide sur la performance d'un apprenant. Au contraire, les compétences de l'apprenant évalué sont, par définition, en construction. C'est pourquoi on pourra considérer qu'il s'agit d'une relation asymétrique et d'un geste évaluatif descendant. Dans la littérature, le terme *hétéro-*

évaluation, employé pour désigner cette modalité évaluative, met l'accent sur cette différence de statut entre les deux acteurs du processus. On retrouve ce terme dans des champs variés comme la psychologie (Lignier *et al.*, 2016), la médecine (Blettery *et al.*, 1996) ou les sciences de l'éducation et de la formation (Bolzan *et al.*, 2022; Chilotti, 2013). Selon (De Ketele, 1993, p. 60), c'est un « acte syncrétique étroitement lié à la personne de l'évaluateur et admis comme tel, dans la mesure où l'évaluateur est celui qui a mené l'apprentissage, revendique être la personne la mieux placée pour bien connaître les performances de ses élèves ou de ses étudiants, prend ses responsabilités avec conscience professionnelle ».

5.2.2 L'auto-évaluation

L'hétéro-évaluation est couramment associée à l'auto-évaluation, c'est-à-dire à une évaluation de l'apprenant par lui-même, selon des critères établis par l'enseignant. La question de la fiabilité est alors posée dans de nombreuses études. Il existe ainsi un important faisceau de preuves suggérant que, pour une évaluation donnée, les notes attribuées par des enseignants (hétéro-évaluation) ou par des apprenants (auto-évaluation) sont globalement proches ; cependant, les élèves les plus fragiles ont tendance à surnoter leur propre travail (Falchikov & Boud, 1989; Mowl & Pain, 1995). En outre, il semblerait que les élèves aient tendance à utiliser une fourchette de notes moins étendue que celle des enseignants (Stefani, 1994).

Dans une méta-analyse, (Dochy *et al.*, 1999) abordent également cette question de la fiabilité ; ils emploient le terme de *précision* (« *accuracy* ») et font notamment référence, sur ce point, à deux études : l'une conduite par (Longhurst & Norton, 1997) et l'autre conduite par (Zoller & Ben-Chaim, 1997).

(Longhurst & Norton, 1997) ont conduit leur étude auprès de 67 étudiants en psychologie. Les auteurs ont mis au jour une corrélation positive entre les hétéro-évaluations et les auto-évaluations des productions écrites des étudiants ($r = 0.43$). Les résultats ont également indiqué que, dans l'ensemble, les étudiants faisaient preuve de précision pour évaluer au moyen d'une note leur propre production écrite, mais qu'ils étaient moins précis lorsqu'il s'agissait d'employer une approche par compétences. D'autre part, les étudiants les moins motivés et les plus en difficulté semblaient s'être moins bien approprié la signification des critères issus de cette approche par compétences.

(Zoller & Ben-Chaim, 1997) ont quant à eux étudié la capacité à s'autoévaluer de 71 étudiants en biologie. Leurs résultats ont mis en évidence une différence significative entre les auto-évaluations des étudiants et les hétéro-évaluations des enseignants. Comme facteur

explicatif, les auteurs évoquent un manque d'articulation entre les activités d'apprentissage et les activités d'évaluation dans l'enseignement des sciences à l'époque où leur étude fut conduite.

Les travaux d'autres auteurs nous incitent cependant à relativiser le poids de ce critère de fiabilité lorsqu'il s'agit d'auto-évaluation. C'est notamment le cas de (Topping, 2003), qui précise que beaucoup de ces études portent en fait sur la *validité* de l'auto-évaluation, et non pas sur sa fiabilité :

« On compare des auto- [ou des inter-] évaluations avec des évaluations effectuées par des professionnels, plutôt qu'avec celles effectuées par d'autres élèves ou par les mêmes élèves au fil du temps. »

(Topping, 1998, p. 257)

De surcroît, (Devenney, 1989) a souligné que la fonction de l'évaluation par l'enseignant tend à différer de celle de l'évaluation par l'élève : ainsi, la comparaison entre les deux n'est pas nécessairement pertinente, et une faible fiabilité de l'auto-évaluation pourrait ne pas être rédhibitoire.

Surtout, l'auto-évaluation semble propice à une meilleure conscientisation par les apprenants des efforts à fournir en vue de maîtriser les compétences ciblées, ainsi qu'à une amélioration des acquisitions notionnelles, ce qui tend à renforcer l'aspect formatif de cette modalité évaluative (Thobois-Jacob & Chevry-Pébayle, 2020), avec potentiellement un effet significatif sur les performances des apprenants (Thobois-Jacob *et al.*, 2020) – à condition toutefois de veiller à une bonne appropriation, en amont, des grilles critériées sur lesquelles l'auto-évaluation sera fondée (Thobois-Jacob & Mohib, 2020).

5.2.3 L'inter-évaluation

Si des apprenants s'évaluent mutuellement, on parle alors d'évaluation entre pairs ou d'inter-évaluation. Cette modalité évaluative a fait l'objet de nombreuses publications scientifiques (Chen, 2010; Divjak & Maretić, 2017), dont des méta-analyses riches en enseignements, comme celle de (Topping, 1998). Il en ressort plusieurs points essentiels, notamment le fait que l'inter-évaluation bénéficie à la fois au pair évaluateur et au pair évalué, en particulier à la faveur des compétences métacognitives qui sont mobilisées de part et d'autre. D'autre part, des inhibitions socioaffectives et des biais liés à l'absence d'anonymat peuvent intervenir en situation d'inter-évaluation traditionnelle/manuscrite (Dochy *et al.*, 1999; Topping, 1998). Dans ce contexte, le recours au numérique pourrait permettre d'anonymiser

les travaux des élèves, potentiellement en double aveugle – *i.e.* pour que l'élève évaluateur ne sache pas qui est l'auteur du contenu qu'il évalue ; tandis que l'élève évalué ignorerait lui aussi l'identité du pair évaluateur.

5.3 La coévaluation

5.3.1 Une quatrième modalité hybride

De fait, il existe une quatrième modalité évaluative : la coévaluation. Apparue plus tardivement, d'abord dans le champ de la formation puis dans celui de l'éducation, elle présente vis-à-vis des trois modalités évaluatives précédemment abordées une différence intrinsèque, en raison de son caractère hybride. Cette hybridité provient d'une multimodalité évaluative, *i.e.* un mélange d'hétéro- et d'auto-évaluation, ou d'hétéro- et d'inter-évaluation. Certains dispositifs d'évaluation vont même jusqu'à mobiliser, pour une activité donnée, les trois modalités à la fois : hétéro-évaluation, auto-évaluation et inter-évaluation (Arroyave Buriticá *et al.*, 2007). Cette multimodalité peut servir de point d'appui à un dialogue entre apprenants, ou entre enseignant et apprenant(s) (Dochy *et al.*, 1999), ce qui contribue à rendre l'évaluation coopérative.

5.3.2 Un concept polysémique

La coévaluation a longtemps été peu documentée, probablement en raison des difficultés techniques liées à sa faisabilité (Segers *et al.*, 2003a). À partir des années 1990-2000, le développement de plus en plus large des applications numériques semble avoir rendu ces approches hybrides plus accessibles (Braun, 2003). Un terme souvent employé aujourd'hui pour désigner cette évaluation multimodale est effectivement celui de *coévaluation* (Dochy *et al.*, 1999). Originellement, ce terme a émergé dans le contexte de la formation initiale des enseignants. Il désignait des situations mélangeant auto-évaluation par l'enseignant stagiaire et hétéro-évaluation par le formateur (Colognesi & Van Nieuwenhoven, 2017). À la faveur d'un élargissement du sens, le terme coévaluation peut aujourd'hui également s'employer en contexte d'enseignement/apprentissage scolaire pour mettre l'accent sur le partage du geste évaluatif entre différents acteurs de statut différent : l'enseignant, le pair, voire soi-même. Le terme peut ainsi renvoyer à différentes combinaisons évaluatives, ce qui entretient de facto un certain flou terminologique.

5.3.3 Coévaluation et différenciation pédagogique

Si la coévaluation peut désigner diverses combinaisons d'hétéro-, inter- et auto-évaluation, (Quesada *et al.*, 2019) insistent quant à eux sur la participation des apprenants au processus évaluatif, un processus négocié entre enseignants et apprenants dans le but d'atteindre un consensus : la responsabilité du geste évaluatif est partagée et repose sur le dialogue. Ce processus de dialogue et de négociation permet aux enseignants d'obtenir des informations pertinentes concernant les apprentissages, tout en fournissant aux apprenants des possibilités accrues en termes d'apprentissage et d'autorégulation. Cette approche de l'évaluation relève d'un paradigme collaboratif : l'enseignant consent à une forme de décentration, la relation avec les apprenants devenant moins verticale et descendante, et davantage horizontale et ascendante (Allal, 2020). Pour autant, dans cette conception de la coévaluation, l'enseignant peut conserver une part de contrôle sur le processus évaluatif pour en garantir la validité (Topping, 1998).

Ainsi conçue, la coévaluation peut favoriser l'autorégulation des apprentissages (Hawe & Dixon, 2017; Meusen-Beekman *et al.*, 2016; Nicol & Macfarlane-Dick, 2006), avec des effets significatifs en termes de motivation (Dignath *et al.*, 2008), de métacognition (Deeley, 2014; Sivan, 2000) et de sentiment d'efficacité personnelle (Boud & Falchikov, 2006; Orsmond *et al.*, 2013) – bref, d'autodirection des apprentissages au sens de (Carré, 2010). De même, (Hortigüela Alcalá *et al.*, 2019) ont montré dans une étude à grande échelle (N = 1021) portant sur une population d'étudiants à l'université, que les pratiques coévaluatives permettent une hausse significative des compétences transversales.

En outre, dans la mesure où les apprenants sont associés au processus évaluatif (Dochy *et al.*, 1999), la coévaluation peut contribuer à la mise en œuvre des pédagogies coopératives en favorisant la constitution de ce que (Lescouarch, 2021) appelle un *collectif apprenant*. En effet, la coévaluation promeut une approche plus démocratique de l'éducation car elle contribue à empuissanter les élèves (« *empowerment* ») : elle implique une distribution – même partielle – des enjeux du processus évaluatif entre l'enseignant et les apprenants (Blau & Shamir-Inbal, 2018; Connac & Cieutat, 2021; Quesada *et al.*, 2019; Reynaud, 2022). L'ensemble de ces éléments contribue à faire de la coévaluation un levier efficace pour mettre la différenciation pédagogique en œuvre (Allal, 2017).

5.3.4 La coévaluation individualisée

Pour que la coévaluation permettent aux apprenants aux élèves de découvrir et d'évaluer les travaux *individuels* de leurs pairs, et d'exercer leurs compétences métacognitives *individuelles*, un obstacle demeure cependant.

En effet, nous avons vu plus haut que l'évaluation entre pairs, et par extension, la coévaluation, posent des problèmes de biais sociocognitifs et d'inhibitions socioaffectives pour les élèves – qu'ils soient en position d'évaluateur ou d'évalué. Une anonymisation en double aveugle permettrait de solutionner ce problème. Néanmoins, bien qu'une telle anonymisation soit techniquement réalisable de manière manuscrite, sa généralisation à l'échelle d'un groupe classe de taille moyenne, travaillant individuellement, constituerait un processus fastidieux et chronophage, donc difficilement répliquable. Or le manque de continuité pose problème pour la mise en œuvre de ce type de pratique évaluative (Quesada *et al.*, 2019).

Certes, pour contourner cette difficulté, la coévaluation est usuellement pratiquée par petits groupes. Cette approche groupale ne nécessite pas d'anonymiser les travaux dans la mesure où, eu égard à la production et à l'inter-évaluation d'une production écrite, la dilution de la responsabilité individuelle au sein d'un collectif permet de limiter les biais sociocognitifs et les inhibitions socioaffectives. Il n'en demeure pas moins que cette coévaluation par petits groupes ne permet pas aux apprenants d'évaluer les travaux *individuels* de leurs pairs, et d'exercer leurs compétences métacognitives *individuelles* en confrontant leur jugement personnel à celui de l'expert qu'est l'enseignant.

Ainsi, pour pratiquer la coévaluation formative individualisée dans de bonnes conditions – *i.e.* avec la mise en œuvre d'une anonymisation en double aveugle (auteur/évaluateur) –, il s'avère nécessaire de recourir à un artéfact numérique, ce qui fait l'objet du prochain chapitre.

Ce qu'il faut retenir

Depuis la fin du XXe siècle, le passage progressif de l'ère du test à l'ère de l'évaluation a permis de mieux discerner les quatre finalités évaluatives – diagnostique, formative, sommative et certificative – et leurs rôles complémentaires au service de la réussite de tous les élèves. Si l'évaluation est la plus souvent l'apanage de l'enseignant (hétéro-évaluation), d'autres modalités sont devenues de plus en plus courantes et de mieux en mieux documentées dans la littérature scientifique, comme l'auto-évaluation ou l'inter-évaluation.

Une quatrième modalité, hybride, a également émergé : la coévaluation, qui peut constituer un levier particulièrement efficace pour mettre en œuvre la différenciation pédagogique ; à fortiori si elle est envisagée de manière individualisée. Cela pose toutefois le problème des biais sociocognitifs, que seule une anonymisation en double aveugle permet de solutionner. Or une telle anonymisation nécessite de recourir au numérique.

6 La coévaluation numérique individualisée

Avant d'envisager des ressources numériques qui permettraient de pratiquer la coévaluation au collègue – notamment avec anonymisation en double aveugle –, il s'agit de définir un cadre théorique qui permettra d'en conceptualiser les usages – et éventuellement les mésusages, voire les non-usages. Nous emprunterons pour ce faire au champ de l'ergonomie, en nous axant sur les usages des technologies numériques en éducation et en formation.

6.1 La genèse instrumentale

La *genèse instrumentale* s'inscrit dans le cadre de la théorie instrumentale de (Rabardel, 1995) et désigne le processus par lequel un *artéfact* devient *instrument*. Chez cet auteur, la notion d'artéfact renvoie à la fois aux objets culturels ou symboliques – *i.e.* aux connaissances – et aux « objets matériels construits, que sont les outils fabriqués » ; et c'est seulement « lorsqu'ils entrent en interaction avec le sujet » que des artéfacts peuvent devenir des instruments (Marquet, 2005b, p. 115). Ce processus s'accomplit par le truchement de *schèmes d'utilisation*, *i.e.* des virtualités d'usage qui sont actualisées, soit parce que le sujet s'adapte à l'artéfact (*instrumentation*), soit parce que l'artéfact est adapté par le sujet à ses propres besoins (*instrumentalisation*). Lorsque l'instrumentalisation va jusqu'à détourner l'usage initial de l'artéfact, on parle de *catachrèse*. La Figure 5 récapitule cette genèse instrumentale.

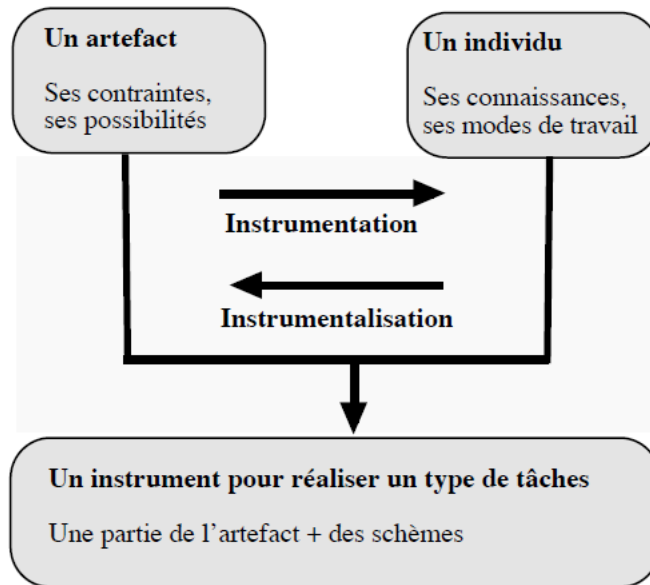


Figure 5 : La genèse instrumentale, combinaison de deux processus (Trouche, 2007)

Dans le contexte des nouvelles technologies de l'information et de la communication, encore émergeant à l'époque de la rédaction de son ouvrage phare sur la genèse instrumentale, Rabardel précisait déjà que « chaque acteur détournant et réinterprétant les possibilités d'usage d'une technologie intellectuelle, lui confère, de ce fait, un nouveau sens » (Rabardel, 1995, p. 130), préfigurant par là-même que des catachrèses sont susceptibles de survenir fréquemment pour les utilisateurs d'environnements informatisés pour l'apprentissage humain (EIAH). L'auteur ajoute que « les transformations matérielles de l'artefact [...] ne sont d'ailleurs pas limitées à la structure de l'artefact, et peuvent aussi concerner le plan du fonctionnement, y compris pour des artefacts relevant des technologies informatiques » (Rabardel, 1995, p. 115).

6.2 Le conflit instrumental en EIAH

Dans le prolongement de Rabardel, (Marquet, 2005a) souligne la pertinence du concept de conflit instrumental dans le contexte des environnements informatisés pour l'apprentissage humain et en approfondit l'analyse.

6.2.1 Triple enchâssement des artefacts

L'auteur met notamment en évidence un *triple enchâssement* entre les artefacts didactique, pédagogique et technique en EIAH (Figure 6) :

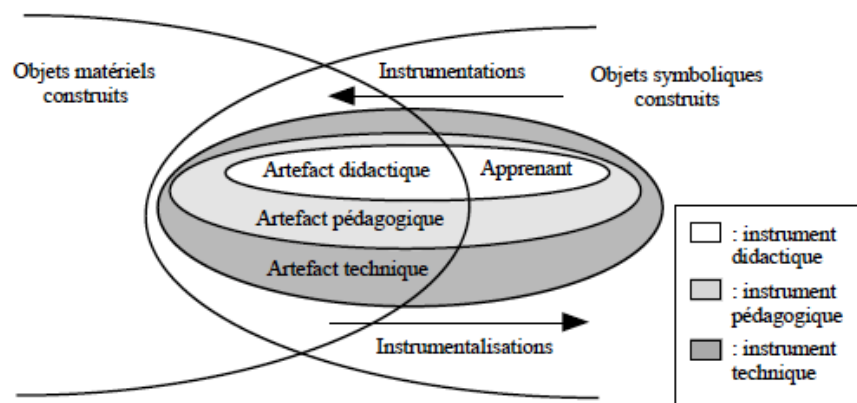


Figure 6 : Les niveaux d'emboîtement des instruments (Marquet, 2005a)

La figure ci-dessus montre en effet que la genèse instrumentale s'effectue à plusieurs niveaux, et comment ces dimensions sont enchâssées les unes dans les autres. À un niveau fondamental, on trouve l'artéfact didactique, c'est-à-dire les contenus (savoirs, savoir-faire et/ou savoir-être) dont l'enseignant souhaite faire des objets d'apprentissage pour l'apprenant.

Ensuite se pose la question du « comment ? » : comment ces contenus pourront-ils être enseignés, avec quelle approche pédagogique ? C'est ce que l'on nomme ici l'artéfact pédagogique. Ce dernier englobe l'artéfact didactique, dans la mesure où la dimension pédagogique est en quelque sorte le véhicule conditionnant la manière dont l'artéfact didactique sera délivré aux élèves.

Enfin, l'EIAH implique un emboîtement à un dernier niveau, qui englobe les deux autres : il s'agit de l'artéfact technique, moyen matériel grâce auquel les deux premiers artéfacts enchâssés pourront parvenir à leur but. Il s'agit ici de la concrétisation des objectifs d'apprentissage visés. Dans le cas des EIAH, cet artéfact technique/numérique sera lui-même souvent de nature composite : à la fois matériel (« *hardware* ») et logiciel (« *software* »).

Ce triple enchâssement des artéfacts est également conceptualisé et opérationnalisé par (Boris *et al.*, 2013; Koehler & Mishra, 2005, 2009) dans des études empiriques qui montrent notamment que les pédagogies actives – *i.e.* des pédagogies qui promeuvent l'autodirection des apprentissages et la coopération – permettent aux apprenants de développer une meilleure compréhension du réseau complexe de relations enchâssées qu'entretiennent les contenus didactiques, les approches pédagogiques et les technologies numériques.

Par ailleurs, outre la notion d'artéfact, d'autres concepts précédemment cités et issus de Rabardel (1995) comme l'instrumentation, l'instrumentalisation, la catachrèse (*i.e.* le détournement d'usage) peuvent s'appliquer avec pertinence en EIAH. En particulier, le concept de *conflit instrumental* y revêt un intérêt tout particulier (Marquet, 2005a). Ce concept

correspond à une situation dans laquelle « l'écran fait écran à la connaissance » (Marquet, 2011, p. 37). En effet, selon (Marquet, 2005b, p.116) :

« De même que certains procédés pédagogiques ne conviennent pas à certains contenus disciplinaires, ou que certains contenus disciplinaires ne s'accommodent pas de certaines approches pédagogiques, l'ajout d'un artefact technique peut provoquer des déséquilibres. »

En d'autres termes, lorsque les apprentissages sont partiellement ou totalement empêchés en EIAH, le concept de conflit instrumental consiste à postuler qu'il y a « incompatibilité des objets didactiques et pédagogiques qui sont associés avec l'objet technique qui est introduit » (Marquet, 2011, p. 39).

Pour commencer, il peut concerner deux types d'utilisateurs : les enseignants d'une part, et les élèves d'autre part. De plus, trois types de situations sont à envisager :

- En amont de la situation d'enseignement/apprentissage
- Pendant la situation d'enseignement/apprentissage
- En aval de la situation d'enseignement/apprentissage

6.2.2 Les trois degrés de conflictualité

Dans le prolongement des travaux de (Zeller & Marquet, 2021a, 2021b) , le concept de conflit instrumental permet en outre de dénombrer trois types d'observables en EIAH, classés ici par ordre croissant de conflictualité. Lorsqu'en raison d'une difficulté rencontrée dans l'instrumentation ou l'instrumentalisation d'un artefact technique – *i.e.* face à un conflit instrumental –, un élève ou un enseignant :

- Est retardé dans l'accomplissement d'une tâche, on parlera de *retard* (d'apprentissage ou d'enseignement)
- Doit renoncer à l'accomplissement d'une partie d'une tâche, on parlera d'évitement, ou de *renoncement partiel* (à une activité d'apprentissage ou à une situation d'enseignement)
- Doit renoncer à l'accomplissement de la totalité d'une tâche, on parlera de *renoncement total* (à une activité d'apprentissage ou à une situation d'enseignement).

6.2.3 Le concept de régulation instrumentale

Dans la typologie précitée, les notions de *retard* et de *renoncement partiel* permettent d'envisager des situations d'enseignement/apprentissage en EIAH dans lesquelles le conflit instrumental ne constitue pas un obstacle insurmontable pour l'utilisateur. De fait, selon (Zeller & Marquet, 2021b, p. 9), les usages du numérique pour l'éducation « génèrent des conflits instrumentaux tout en permettant paradoxalement de les gérer, dans un processus d'ajustement ».

De même que le conflit instrumental est selon (Marquet, 2005b, p. 116) « la manifestation d'interférences entre les artefacts didactique, pédagogique [et] technique », ces ajustements, ou *régulations instrumentales*, peuvent porter sur un ou plusieurs de ces trois artefacts. Ces stratégies ad hoc, improvisées par l'utilisateur en réponse à une difficulté d'instrumentation ou d'instrumentalisation, restent cependant à investiguer (Zeller & Marquet, 2021b).

6.3 Artéfacts numériques pour la coévaluation

6.3.1 Une pratique encore émergente

Si l'*inter-évaluation* numérique est bien documentée dans la littérature (Bouzidi & Jaillet, 2009; Lu & Law, 2012; Millard *et al.*, 2008; Wen & Tsai, 2006; Yang & Tsai, 2010), on trouve peu de travaux concernant la *coévaluation* numérique en milieu scolaire (Mostert & Snowball, 2013). (Kaplan, 2019) analyse en profondeur les stratégies (méta-)cognitives susceptibles de s'y déployer, mais sans donner de détails concernant les plateformes d'*e-learning* utilisées pour coévaluer dans son étude. (Fauré *et al.*, 2018) soulignent les opportunités liées au numérique pour la mise en œuvre de pratiques coévaluatives, mais leur étude se situe dans le cadre de la formation des enseignants et ne formule pas de préconisations spécifiques quant au choix de tel ou tel type de ressource numérique. De même, les travaux de (Tali & Germier, 2021) concernent l'utilisation d'un tableau de bord en ligne collaboratif qui peut servir de point d'appui à des pratiques coévaluatives, mais le modèle qu'ils proposent concerne la formation des enseignants. L'étude menée par (Blau & Shamir-Inbal, 2018) consistait quant à elle en l'utilisation d'une variété de ressources privatives en ligne dédiées au travail collaboratif en général (*e.g.* traitement de texte, tableur, visioconférence) sans recourir à une application numérique spécifique pour la coévaluation proprement dite. Enfin, à travers une application numérique accessible via smartphone, conçue spécialement pour pratiquer la

coévaluation selon une approche par compétences, (Ishikawa *et al.*, 2021) ont montré une augmentation significative de la motivation chez les étudiants participants à l'étude.

6.3.2 Solutions privatives payantes

De nos jours, pour pratiquer l'*inter-évaluation* numérique, on peut trouver en ligne des solutions logicielles privatives et payantes, comme peerassessment.com ou peerstudio.org ; ainsi que des extensions applicatives gratuites mais fonctionnant sur des plateformes payantes développées par des GAFAM, *e.g.* pour Google Classroom ou Microsoft Teams. Toutefois, le recours à de telles solutions privatives peut poser un problème en termes de protection des données des élèves et de respect du règlement général sur la protection des données (RGPD). Par ailleurs, cela représente un investissement budgétaire significatif, qui peut poser problème en contexte scolaire, notamment au collège ; d'autant plus que depuis une dizaine d'années les collèges français disposent tous déjà, via leur environnement numérique de travail (ENT), d'un accès protégé à [Moodle](https://moodle.org) (Figure 7) une plateforme logicielle libre qui permet notamment de pratiquer la *coévaluation* – tandis que les solutions précitées ne permettent que l'*inter-évaluation*.

The screenshot shows the 'Exemple' dashboard for a Moodle Workshop. At the top, there is a 'Marquer comme terminé' button. Below it, the 'Phase de notation des évaluations' section contains a progress table with five columns: 'Phase de mise en place', 'Phase de remise', 'Phase d'évaluation', 'Phase de notation des évaluations', and 'Fermé'. The 'Phase de notation des évaluations' column is highlighted in green and shows 'Phase activée' with a progress indicator. Below the table, there are settings for 'Méthode de notation des évaluations' (set to 'Comparaison avec la meilleure évaluation') and 'Réglages de la notation des évaluations' (set to 'Normal'). A 'Recalculer les notes' button is visible. At the bottom, there is a 'Rapport d'évaluation de l'atelier' link.

Figure 7 – Exemple de « tableau de bord » d'un Atelier Moodle

6.3.3 Inter-évaluer avec Moodle Atelier

(Marín & Pérez Garcias, 2016) ont fait usage de cette plateforme pour pratiquer l'*inter-évaluation*, en mobilisant un module dédié, intitulé *Moodle Atelier* (« *Moodle Workshop* »). (Mostert & Snowball, 2013) ont notamment montré que le recours à Moodle Atelier facilitait significativement l'*inter-évaluation*, et permettait d'accéder à des fonctionnalités absentes en

contexte manuscrit, *e.g.* l'anonymisation en double aveugle. (Bouziane & Ziad, 2018) ont quant à eux montré que la coévaluation pratiquée avec Moodle Atelier permettait d'améliorer significativement la métacognition des élèves participant à l'étude, en l'occurrence la qualité/pertinence des feedbacks entre élèves. Pour leur part, (Llera *et al.*, 2015) ont pratiqué avec Moodle Atelier une forme de *coévaluation* consistant, en amont, à associer les élèves à l'élaboration des critères de réussite qui servaient ensuite à l'inter-évaluation.

6.3.4 Une catachrèse pour coévaluer avec Moodle Atelier

Porté par une communauté vivante et dynamique de développeurs et d'utilisateurs, le logiciel libre Moodle offre l'avantage d'être particulièrement flexible en termes de paramétrage – donc propice aux instrumentalisation, voire aux *catachrèses*, au sens de (Rabardel, 1995). Ainsi, bien qu'initialement prévu pour l'*inter-évaluation*, l'usage de Moodle Atelier peut être détourné, au moyen d'opérations de paramétrage avancé, afin de pratiquer la coévaluation. Précisons qu'il s'agit bien de moyens détournés : le paramétrage de Moodle Atelier ne prévoit tout simplement pas la coévaluation ; toutefois, sa programmation délibérément ouverte, typique du logiciel libre et ainsi voulue par ses développeurs, permet de mobiliser un nombre virtuellement infini de combinaisons de paramètres avancés, susceptibles d'altérer significativement le mode de fonctionnement du logiciel.

En l'occurrence, il s'agit d'appliquer à l'artéfact technique Moodle Atelier des *schèmes d'utilisation* (Rabardel, 1995) non initialement prévus – ceux de la coévaluation –, au moyen d'une combinaison contre-intuitive de paramètres avancés. Cette combinaison n'existe dans Moodle Atelier qu'à l'état de virtualité latente. Le détournement d'usage nécessaire pour actualiser ce potentiel – et permettre effectivement la coévaluation, au sens d'un authentique partage du geste évaluatif entre enseignant et apprenant (Quesada *et al.*, 2019) – relève de ce que (Rabardel, 1995) nomme une *catachrèse*. Celle-ci est le fruit de tâtonnements expérimentaux, auxquels nous avons procédé lors de la phase exploratoire de la présente recherche. Précisons que les paramètres avancés que nous avons mobilisés pour ce faire sont relativement difficiles d'accès, et susceptibles de compromettre le fonctionnement du logiciel en cas d'erreur. Au final, la catachrèse peut se résumer aux trois opérations détaillées ci-après.

En premier lieu, dans la section « Permissions > Activité : Atelier » de l'atelier Moodle, la suppression locale de la permission « Voir les noms des auteurs » pour les « Etudiants » est une opération d'instrumentalisation basique, permettant d'accommoder l'artéfact technique au schème d'utilisation « inter-évaluation en *double aveugle* » (Figure 8).

Figure 8 – Schème d'utilisation « inter-évaluation en double aveugle »

Bien que l'opération d'instrumentalisation susmentionnée constitue un préalable indispensable, c'est seulement en la combinant aux deux opérations suivantes qu'on aboutit à la catachrèse recherchée. Ces deux opérations sont indissociables dans la mesure où elles se contrebalancent mutuellement pour produire l'effet recherché.

Tout d'abord, dans « Rôles attribués localement », il s'agit d'attribuer artificiellement à l'enseignant le rôle local d'« Etudiant¹⁶ » (Figure 9).

Attribuer le rôle « Étudiant » dans Atelier: Exemple

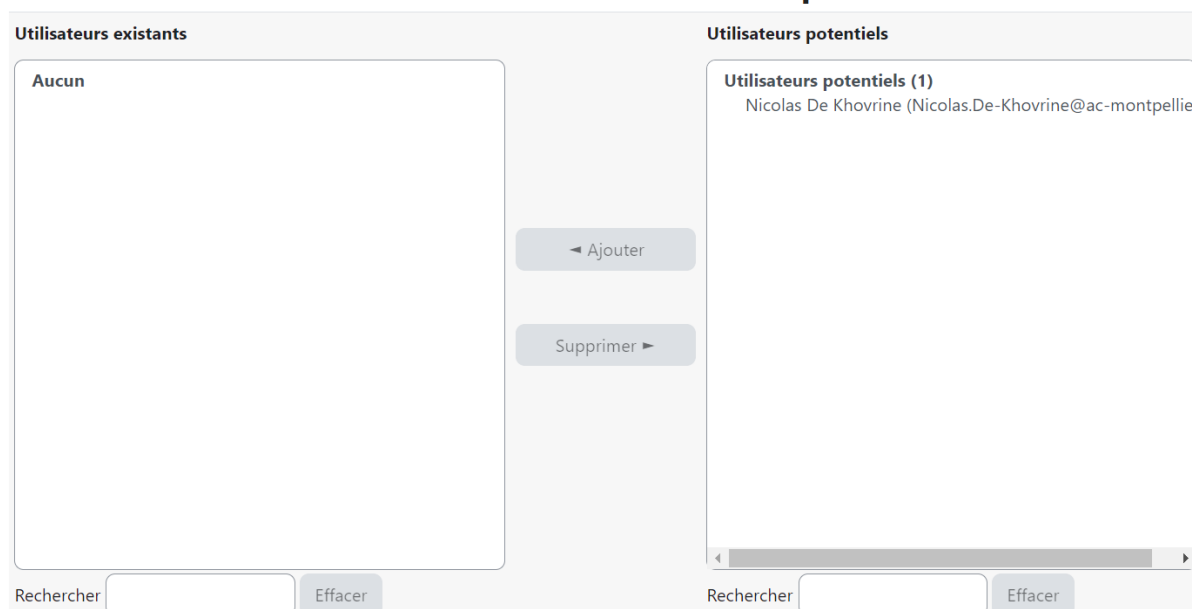


Figure 9 – Attribuer le rôle « Etudiant » à l'enseignant

En complément, il s'agit de permettre au logiciel de reconnaître malgré tout que l'enseignant, ainsi *déguisé* en « Etudiant », est bel et bien un enseignant – *i.e.* celui dont l'hétéro-évaluation servira d'étalon de mesure. En bref, le fait de pondérer l'hétéroévaluation de l'enseignant au moyen du coefficient maximal – *i.e.* coefficient 16 – garantit cela (Figure 10). Pour mieux comprendre l'importance de cette seconde facette de la catachrèse *coévaluation* dans Moodle Atelier, il convient d'explicitier la manière dont ce logiciel calcule ses scores.

¹⁶ Le vocable « Etudiant » est ici un terme générique correspondant au langage utilisé dans Moodle pour différencier les différents types d'utilisateurs d'un cours ou d'une activité, e.g. « Enseignant éditeur », « Enseignant non-éditeur », « Etudiant » ou « Gestionnaire ».

Modalités d'évaluation

- Pondération de l'évaluation: 16
- Note calculée pour l'évaluation: 15
- Modifier la note de l'évaluation: 14
- Feedback pour l'évaluateur: 13

Other options in the dropdown: 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0.

Figure 10 – Attribuer la pondération 16 à l'hétéro-évaluation de l'enseignant « déguisé » en « Etudiant »

En effet, Moodle Atelier calcule au final *deux types de scores*, *i.e.* la « note pour le travail remis » et la « note pour l'évaluation ». Ces deux types de scores sont des données quantitatives calculées par le logiciel à partir de données initialement qualitatives, en l'occurrence les coévaluations réalisées par l'enseignant et les apprenants selon une approche par compétences, *i.e.* de manière non chiffrée, sur une échelle de Likert à cinq points allant de « Insuffisant » à « Très satisfaisant » (Figure 11).

Formulaire d'évaluation

Grille

Critères	Niveaux			
Le récit contient un dialogue au discours direct, correctement ponctué et disposé.	<input type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
Le lieu sera décrit.	<input type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
Deux des 5 sens seront mis en évidence.	<input type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
Le texte comportera des indications sur le physique et sur le caractère d'un des personnages au moins.	<input type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
Des éléments de la culture japonaise seront perceptibles dans le texte.	<input type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant

Figure 11 – Formulaire d'évaluation avec échelle de Likert à cinq points

Une fois calculées par le logiciel, les « notes pour le travail remis » sont comparées entre elles pour générer le second type de score, *i.e.* la « note pour l'évaluation ». Or sur le site [Moodle.org](https://moodle.org), il est précisé que le mode de calcul des notes pour l'évaluation est « déterministe ». Autrement dit, l'algorithme y a pour but de déterminer une note « consensuelle », au sens statistique du terme :

La note la plus proche de la moyenne des notes est considérée comme étant « la meilleure » par l'algorithme. On suppose que cette note serait consensuelle pour la majorité des examinateurs.

A cette note de référence est attribué un poids de 100%.

A chacune des autres notes, on attribue un poids inférieur à 100%, en proportion de l'écart (différence) constaté par rapport à la note de référence (« la meilleure »)¹⁷.

Or pour pratiquer la *coévaluation* dans Moodle Atelier de manière *valide* au sens de (Topping, 2003), la « note de coévaluateur » doit rendre compte de la pertinence de la coévaluation de l'élève, non pas en prenant le « consensus » des autres élèves coévaluateurs comme référence – car ce consensus pourrait aboutir à une sous-évaluation ou à une sur-évaluation (Falchikov & Boud, 1989; Longhurst & Norton, 1997; Mowl & Pain, 1995; Stefani,

¹⁷https://docs.moodle.org/4x/fr/Utiliser_Atelier#Estimation_de_l.27.C3.A9valuation_des_pairs

(consulté le 8/12/2022)

1994) –, mais en prenant l'*hétéro-évaluation* de l'enseignant comme référence. On affecte donc à cette dernière la pondération maximale possible dans Moodle Atelier (*i.e.* coefficient 16), tandis qu'on affecte aux coévaluations d'élèves la pondération par défaut (*i.e.* coefficient 1). En limitant par ailleurs le nombre de coévaluations d'élèves à 4 (Figure 12), on garantit que, même en cas de consensus erroné de leur part, l'hétéro-évaluation de l'enseignant conservera un poids statistique quatre fois plus important que l'ensemble des coévaluations d'élèves. De sorte que, lors du calcul par Moodle Atelier de la moyenne des « notes pour le travail remis » (élèves et enseignant confondus), l'algorithme déterministe considère obligatoirement l'hétéro-évaluation de l'enseignant comme la « note de référence » – *i.e.* la coévaluation à l'aune de laquelle la pertinence/proximité des coévaluations d'élèves est mesurée. Précisons qu'avec Moodle Atelier cette comparaison ne s'effectue pas simplement en bout de chaîne, en prenant en considération la seule « note pour le travail remis » (calculée par le logiciel à partir des compétences via l'échelle de Likert) : l'algorithme y effectue cette comparaison de manière fine, *i.e.* compétence par compétence – ou « critère par critère », pour reprendre la terminologie propre à Moodle Atelier.

Évaluation par les pairs

Atelier Paramètres Formulaire d'évaluation Attribution des travaux remis Plus ▾

Attribution aléatoire ▾

Attribution aléatoire

▾ Réglages des attributions aléatoires

Mode de groupe Aucun groupe

Nombre d'évaluations par évaluateur ▾

Supprimer les attributions actuelles

Les participants peuvent évaluer sans avoir remis de travail

Ajouter les auto-évaluations Auto-évaluation désactivée

Enregistrer

Figure 12 – Limitation à 4 coévaluations par élève lors de l'attribution aléatoire des travaux

Ajoutons pour finir que les coévaluations des élèves conservent un impact sur la « note pour le travail remis ». Cet impact est marginal compte tenu de la différence de pondération

explicitée plus haut ; pour autant il n'est pas totalement négligeable, puisque le coefficient affecté aux coévaluations d'élèves est 1 – et non pas 0. D'autre part, concernant la « note de coévaluateur », le consensus est obtenu de manière entièrement automatisée, *i.e.* via une prise en charge par l'algorithme de l'artéfact numérique. Enfin, les écrits commentatifs qui accompagnent les coévaluations peuvent servir de point de départ efficace pour instaurer un dialogue et une éventuelle négociation entre l'enseignant et les apprenants relativement aux notes données ou reçues – à fortiori si la finalité de la coévaluation est formative. En somme, dans le cadre de la catachrèse élaborée pour la présente recherche, l'usage de Moodle Atelier est effectivement détourné pour calculer deux types de notes qui sont issues, non pas de l'inter-évaluation, mais de la *coévaluation* au sens de (Quesada *et al.*, 2019, p. 988, traduction libre) :

« La note finale est le fruit d'une élaboration conjointe de la part de l'enseignant et de l'apprenant, au moyen du dialogue et de la négociation ».

6.3.5 La fonctionnalité « Travaux exemplaires »

Par l'entremise d'une fonctionnalité nommée « Travaux exemplaires », Moodle Atelier permet de mettre en œuvre, en amont de l'activité de coévaluation proprement dite, des entraînements destinés à familiariser l'apprenant avec cette activité. Concrètement :

« Si ce réglage est activé, les participants pourront évaluer un ou plusieurs travaux exemplaires et comparer leur évaluation avec celle fournie en référence. La note n'est pas comptée dans la note d'évaluation »¹⁸.

Ainsi, ces scénarios interactifs sont mis en œuvre par l'enseignant dans l'artéfact numérique, par l'entremise d'« exemples de travaux », qui constituent en quelque sorte des exercices d'appropriation, voire des évaluations formatives pour lesquelles les élèves bénéficient d'un feedback personnalisé, automatisé et surtout, immédiat. De surcroît, comme dans la coévaluation à laquelle il prépare, ce feedback automatisé et immédiat reflète fidèlement le degré de maîtrise des (méta)compétences cibles manifesté par l'élève, à la faveur de l'algorithme grâce auquel l'artéfact numérique effectue une comparaison pair-à-pair, compétence par compétence, entre l'inter-évaluation de l'élève et l'hétéro-évaluation de l'enseignant (*i.e.* l'évaluation de référence).

¹⁸ Source : Moodle.org, consulté le 18/06/2023 à l'adresse URL : https://docs.moodle.org/4x/fr/Param%C3%A8tres_d'atelier#Travaux_exemplaires.

A dessein, nous ne reprenons pas ici la terminologie propre à la version française de Moodle Atelier, *i.e.* « travaux exemplaires ». En effet cette terminologie peut paraître trompeuse dans la mesure où il ne s'agit pas à proprement parler de travaux « exemplaires » – au sens où ils représenteraient un exemple d'accomplissement idéal de l'activité d'apprentissage, « méritant un 20/20 ». Il s'agit plutôt de « travaux exemples », auxquels les élèves peuvent s'identifier, et qui peuvent mettre en scène des défauts à éviter, selon ce que l'enseignant jugera utile d'y scénariser. Nous nous sommes donc inspirés de la terminologie anglaise originelle (« *example submissions* ») pour proposer l'appellation « Exemples de travaux ».

Ce qu'il faut retenir

Empruntée au champ de l'ergonomie, la théorie de la genèse instrumentale permet d'analyser les usages de la technologie, au moyen de concepts porteurs, *e.g.* artéfact, schèmes d'utilisation, instrumentation, instrumentalisation, catachrèse. Ces concepts peuvent être transposés avec profit dans le champ des usages du numérique pour l'éducation. A ce titre, la notion de conflit instrumental revêt un intérêt particulier pour mieux comprendre et tenter de prévenir les mésusages, voire les non-usages en EIAH, en explicitant notamment le triple enchâssement des artéfacts didactique, pédagogique et technique qui s'y joue. On voit ainsi apparaître une typologie permettant de distinguer trois degrés de conflictualité instrumentale. Ce cadre conceptuel permet alors d'envisager des artéfacts numériques pour pratiquer la coévaluation individualisée en double aveugle. Concernant cette pratique émergente, les solutions privatives payantes ne s'avèrent pas pertinentes, à fortiori dans le contexte du collège ; tandis que le logiciel libre Moodle Atelier y est déjà disponible, et que son usage est bien documenté dans la littérature. Toutefois, il s'agit d'un logiciel conçu pour l'inter-évaluation : une catachrèse est donc nécessaire pour que la coévaluation y devienne possible.

7 Didactisation de la production écrite en EIAH

7.1 Production écrite et étude de la langue au collège

En didactique du français langue première, les approches croisées de la production d'écrit et de l'étude de la langue (Garcia-Debanc, 1993) constituent une entrée pertinente au collège ; en particulier au cycle 4, comme en atteste le curriculum actuellement en vigueur (programmes de 2016, revus en 2018 et 2020). En effet, le programme de français en cycle 4 met particulièrement l'accent sur l'importance d'une articulation entre étude de la langue et production d'écrit :

« Il est important d'établir un lien entre l'écriture de textes et l'étude de la langue, de différentes manières : par anticipation de leçons de langue (écrire avant d'étudier un point de langue), en prolongement (écrire pour appliquer une leçon), en révision (améliorer son écrit grâce au savoir acquis en étude de la langue). La pratique de l'écrit devient ainsi plus réflexive et les élèves acquièrent de l'autonomie dans l'amélioration de leurs écrits. Ils savent utiliser l'écrit pour travailler et apprendre. Ils comprennent qu'un écrit n'est jamais spontanément parfait et qu'il doit être repris pour rechercher la formulation qui convient le mieux, préciser ses intentions et sa pensée, justifier un point de vue. » (Direction générale de l'enseignement scolaire, 2020, p. 18).

7.2 La production écrite sur traitement de texte

Pour didactiser la production d'écrit, le recours au traitement de texte (Doquet-Lacoste, 2006; Piolat & Roussey, 1995) modifie substantiellement le processus d'écriture (Delbrassine, 2018), la posture de l'élève (Piolat & Roussey, 1995) et le repérage de certains ressorts textuels (Mangenot, 2000), notamment pour les élèves en difficulté (Beer-Toker *et al.*, 2009). L'écriture sur traitement de texte a des implications cognitives importantes et tend à reconfigurer les modalités scripturales en faveur de modifications plus fréquentes et de plus grande amplitude apportées au texte source (Joannidès, 2014; Penloup, 2017a; Plane, 1996).

7.3 Production écrite sur traitement de texte et inter-évaluation

Le recours au traitement de texte permet en outre de mettre en œuvre une approche coopérative de l'évaluation, *e.g.* en pratiquant l'inter-évaluation ou la coévaluation (Allal, 2018, 2020; Dekhinet *et al.*, 2008). Il s'agit notamment d'inscrire les apprentissages dans une perspective « située », *i.e.* via des scénarios pédagogiques propices à la contextualisation de contenus didactiques dans le cadre de tâches complexes (Allal, 2001; Brown *et al.*, 1989). De nombreux travaux ont montré les apports de l'évaluation entre pairs pour la didactisation de la production écrite, à la fois en termes de maîtrise des compétences et d'autorégulation des apprentissages (Allal, 2020, 2021; Devenney, 1989; Forrer & Wyant, 2015; Juita & Widiyanto, 2019; K. J. Topping *et al.*, 2000) – et singulièrement en ce qui concerne la métacognition (Allal, 2000; Yarrow & Topping, 2001), à condition toutefois d'être mise en œuvre avec une préparation suffisante en amont (Panadero *et al.*, 2016; Sharan, 2010).

7.4 Dispositifs numériques pour l'inter-évaluation de productions écrites

Certains auteurs comme (Darwanto *et al.*, 2020; Kayacan & Razi, 2017; Strasma, 2009; Warren *et al.*, 2008) ont montré les apports de divers artefacts ou dispositifs numériques pour l'inter-évaluation d'écrits produits sur traitement de texte, notamment Moodle Atelier (Bouziane & Ziad, 2018). Singulièrement, un artefact numérique d'inter-évaluation comme Moodle Atelier permet qu'en fin de tâche les élèves évaluateurs produisent un écrit commentatif à destination de leur pair, *i.e.* un feedback concernant la production écrite inter-évaluée. La didactisation de l'écrit commentatif par l'enseignant présente l'avantage d'articuler compétences lectorales et scripturales (Brunel, 2018; Taous, 2020) de manière engageante pour l'apprenant, à fortiori par le truchement « d'espaces numériques d'échanges au sein des classes » (Taous, 2020 : 1). En l'occurrence, de telles pratiques sont propices à une mise en œuvre selon des modalités hybrides – *i.e.* distancielles et présentes à la fois – et asynchrones, qui facilite le développement métacognitif d'une posture d'auteur impliqué dans la production de son texte et favorise une meilleure maîtrise des compétences d'écriture littéraire (Brunel & Guérin-Callebout, 2018). Dans le prolongement d'auteurs comme (Chapelain, 2017; Lacelle & Lebrun, 2018; Penloup, 2017b), les travaux de (Ahr *et al.*, 2011; Moinard, 2017; Régnard,

2012) montrent que la didactisation de l'écrit commentatif par l'entremise du numérique est une approche en développement. A ce jour toutefois, il demeure difficile de trouver dans la littérature scientifique francophone ou anglophone des études empiriques ayant la coévaluation numérique individualisée de productions écrites pour objet, potentiellement dans la mesure où il s'agit d'une pratique encore émergente.

Ce qu'il faut retenir

La production écrite peut être articulée avec profit à l'étude de la langue, en particulier au cycle 4 du collège, et notamment lorsqu'on la pratique sur traitement de texte. Relativement à la posture et aux processus cognitifs de l'apprenant, la production écrite sur traitement de texte modifie substantiellement le rapport à l'écrit. A cet égard, le recours à l'inter-évaluation implique des effets positifs en termes de maîtrise des compétences, d'autorégulation des apprentissages et de métacognition – notamment parce qu'une telle inter-évaluation peut donner lieu à la production d'écrits commentatifs, *i.e.* des feedbacks entre pairs, propices aux apprentissages. L'inter-évaluation de productions écrites sur traitement de texte peut notamment être instrumentée avec profit au moyen de dispositifs pédagog numériques, et singulièrement grâce à un artefact numérique ad hoc, comme Moodle Atelier. Néanmoins, il est difficile de trouver des études empiriques portant sur la coévaluation numérique de productions écrites, dans la mesure où il s'agit d'une pratique encore émergente.

8 La recherche par méthodes mixtes (RMM)

8.1 Une approche adaptée pour l'étude de la différenciation pédagogique

Selon (Forget, 2017b, p. 50), relativement aux problèmes que pose l'opérationnalisation de la différenciation pédagogique :

« La solution se trouve probablement à l'intersection des approches quantitatives et qualitatives : il s'agirait [...] d'appréhender la réalité en assumant cette complexité tout en se donnant des indicateurs de qualité de différents types. Seule une triangulation des données mettant en dialogue des

effets mais aussi des entretiens avec les acteurs et des observations en contexte pourrait, pensons-nous, réduire le flou qui caractérise encore ce champ dont l'actualité et les potentialités ne posent pas question. »

Ainsi, en complément du paradigme de la recherche collaborative en sciences de l'éducation et de la formation, la recherche par méthodes mixtes (RMM) constituerait une approche épistémologique adaptée à la complexité de la différenciation pédagogique. En raison de son intrication avec l'objet d'étude et de sa spécificité, nous avons choisi de présenter cet aspect méthodologique dans l'état de l'art, mais uniquement de manière théorique. Les choix méthodologiques, opérés pour la présente étude dans le cadre de la RMM, seront présentés *infra*, dans la partie dédiée (partie III).

8.2 Un concept polysémique ?

La RMM est un phénomène relativement ancien en sciences humaines et sociales (Lynd & Lynd, 1959). Le concept demeure cependant polysémique. En comparant 19 définitions recueillies auprès d'experts du domaine, (Johnson *et al.*, 2007, p. 123) en ont proposé une définition synthétique :

« La recherche par méthodes mixtes est un type de recherche dans lequel un chercheur ou une équipe de chercheurs combine des éléments des approches de recherche qualitative et quantitative (e.g., l'utilisation de points de vue qualitatifs et quantitatifs, la collecte de données, l'analyse, les techniques d'inférence) avec l'objectif général d'assurer l'ampleur et la profondeur de la compréhension et de la corroboration. » [traduction libre, (Anadón, 2019, p. 107)]

Par ailleurs, cinq thèmes émergent de l'analyse par comparaison réalisée par (Johnson *et al.*, 2007) :

1. Les types de données susceptibles d'être mélangées ;
2. L'étape de la recherche où le mélange a lieu ;
3. L'ampleur de la recherche avec des méthodes mixtes (RMM) ;
4. Les raisons qui orientent le mélange ;
5. L'orientation de la RMM.

8.3 Stratégies pour l'intégration des données

Le terme *triangulation*, mentionné par (Forget, 2017b, p. 50), peut recouvrir une certaine « confusion terminologique » (Anadón, 2019, p. 112). Dans le champ des RMM, cette notion correspond en fait à une des trois stratégies potentielles pour l'intégration des données qualitatives et quantitatives, à l'instar de la *combinaison* et de la *complémentarité* (Anadón, 2019, pp. 110-111). La triangulation consiste à utiliser deux méthodologies – *i.e.* l'une, qualitative et l'autre, quantitative – pour opérationnaliser un même objet de recherche. Les méthodes sont appliquées de manière indépendante, cependant les résultats ont vocation à être comparés. Avec la combinaison, l'une des deux approches est la méthode principale, tandis que l'autre joue un rôle subsidiaire, et a vocation à renforcer la validité des conclusions premières. Enfin, la complémentarité consiste en l'utilisation de deux approches méthodologiques indépendantes, dans le but d'opérationnaliser deux dimensions distinctes de l'objet de recherche. Les résultats sont analysés séparément.

8.4 Modèles (ou devis) principaux

Dans une perspective plus large, (Creswell & Creswell, 2017) conceptualisent trois modèles (ou devis) principaux pour la RMM. Le modèle parallèle et convergent (Figure 13) consiste à collecter simultanément des données issues d'approches quantitatives et qualitatives et à les analyser séparément, avant d'intégrer les deux types de résultats dans une perspective comparatiste.

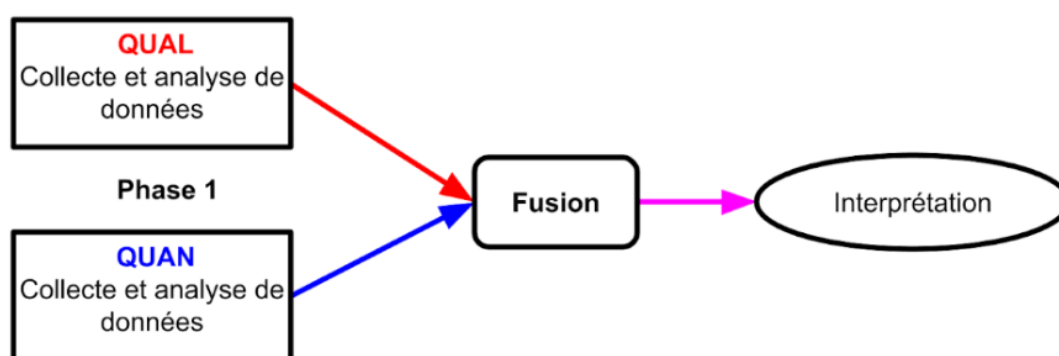


Figure 13 – Modèle parallèle et convergent d'après (Creswell & Creswell, 2017)

Le modèle séquentiel explicatif consiste à utiliser successivement une approche quantitative, puis une approche qualitative dont le but est d'affiner l'interprétation des résultats quantitatifs (Figure 14).

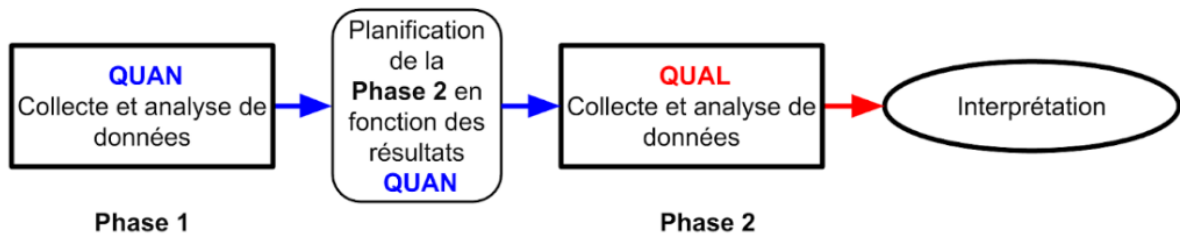


Figure 14 – Modèle séquentiel explicatif

Enfin, le modèle séquentiel exploratoire consiste à utiliser dans un premier temps une approche qualitative (étude exploratoire), dont les résultats servent à mieux cerner les variables nécessaires à l’opérationnalisation d’un objet de recherche, afin d’optimiser l’élaboration du protocole de l’étude quantitative qui suivra (Figure 15).

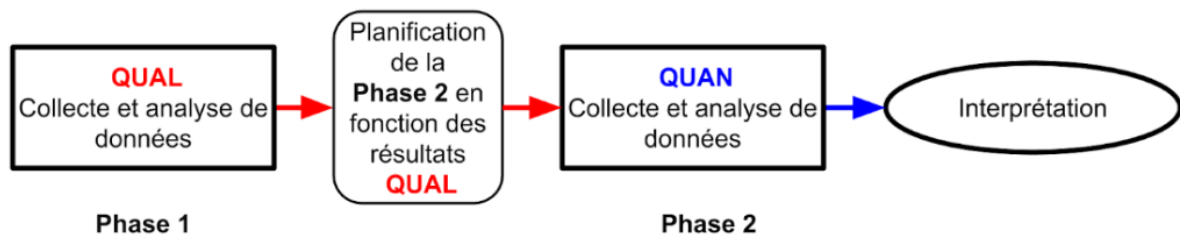


Figure 15 – Modèle séquentiel exploratoire-

Ce qu’il faut retenir

La recherche par méthodes mixtes (RMM) constitue une approche épistémologique pertinente pour tenter d’opérationnaliser la différenciation pédagogique. Elle consiste à mélanger des approches quantitatives et qualitatives. Pour intégrer ces données, on peut recourir à trois stratégies d’intégration différentes : la triangulation, la combinaison et la complémentarité. Essentiellement, la RMM peut par ailleurs être conceptualisée au moyen de trois modèles (ou devis), à savoir : le modèle parallèle et convergent, le modèle séquentiel explicatif et le modèle séquentiel exploratoire.

9 La recherche collaborative

9.1 Une autre approche adaptée à l’étude de la DP ?

Toujours selon (Forget, 2017a, p. 25) :

« Non seulement les propositions [de différenciation pédagogique] se déclinent sous des formes multiples mais chacune d'elles met de surcroît en scène un nombre impressionnant de variables de fonctionnement formant un tout complexe difficilement contrôlable par un plan de recherche expérimentale. [...] Les praticiens ont dès lors une part active à jouer en choisissant, depuis leur contexte, quelles modalités de différenciation expérimenter en classe. Une définition précise des intentions poursuivies par chaque geste/dispositif déployé facilitera l'évaluation des bénéfices visés mais aussi des dérives possibles. ».

Face aux difficultés que soulève l'opérationnalisation de la différenciation pédagogique, on voit ici apparaître d'une part l'importance du rôle des praticiens, et d'autre part la nécessité d'une approche méthodologique adaptée, notamment par le biais de la recherche. Le paradigme de la recherche collaborative en sciences de l'éducation et de la formation (Pentecouteau, 2020) semble donc particulièrement pertinent pour aborder cet objet complexe qu'est la DP. Là encore, bien qu'il s'agisse d'un aspect méthodologique de la présente étude, la spécificité de cette approche épistémologique et son intrication avec notre objet d'étude impliquent que nous en explicitions la dimension théorique dans cet état de l'art – tout en réservant le détail des choix méthodologiques à la partie dédiée (partie III).

9.2 Une posture épistémologique singulière

Dans une recherche-intervention à caractère collaboratif, il s'agit pour le chercheur-intervenant d'expérimenter avec ses partenaires une posture épistémologique permettant de transcender, à l'instar de (Tortérat *et al.*, 2019, 2020), les tensions, catégorisées par (Marcel, 2016a, p. 227) dans sa matrice de la recherche-intervention (Tableau 1), entre « recherche sur » (visée heuristique), « recherche pour » (visée transformative), « recherche avec » (visée participative) et « recherche par » (visée émancipatrice).

Sur	Pour	Avec	Par
Visée heuristique	Visée transformative	Visée participative	Visée émancipatrice
Logique scientifique	Logique professionnelle	Logique collaborative	Logique autonomisante
Connaissance savante	Connaissance pratique et efficace	Connaissance de sens commun	Ouverture du champ des possibles
Validation : procès recherche (critères)	Validation : résolution problème professionnel	Validation : l'inter-reconnaissance	Validation : dynamiques à l'œuvre en dehors et après le dispositif
Universalité des résultats	Contextualisation des résultats	Engagement des acteurs	Co-initiation des acteurs et chercheurs
Troisième monde : le monde de la pensée scientifique (monde culturel)	Premier monde : le monde des objets (monde matériel)	Deuxième monde : le monde des états de conscience, des états mentaux (monde privé)	Troisième monde : le monde de la pensée utopique (monde culturel)
Faire de la recherche	Se servir du potentiel de la recherche	Partager la recherche et l'intervention	Co-renouveler le champ des possibles

Tableau 1 – Matrice de la recherche-intervention (Marcel, 2016a, p. 227)

9.3 Un métissage des savoirs

Au plan épistémique, un dispositif expérimental de coévaluation numérique, ayant la différenciation pédagogique pour objectif, peut devenir ce que (Vinck, 2009, pp. 56-57) nomme un « objet intermédiaire », au sujet duquel (Marcel, 2016b, p. 29) explique qu'il est le fruit d'une « co-élaboration », portée par les différents acteurs d'une recherche collaborative, en appui sur leurs savoirs respectifs. Ces savoirs – celui du chercheur, celui du praticien, celui de l'élève ou du partenaire institutionnel – sont spécifiques et ont leurs langages propres, mais ils ne sont pas irréductibles et sont susceptibles de dialoguer et de se métisser (Marcel, 2010; Rhéaume, 2009).

9.4 Une dimension communautaire

Outre le chercheur-intervenant et les enseignants partenaires, d'autres types d'acteurs peuvent collaborer à une recherche-action collaborative en milieu scolaire, à commencer par les élèves eux-mêmes. Des partenaires institutionnels peuvent également jouer un rôle décisif, *e.g.*, dans l'Éducation nationale : la Cellule Académique Recherche Développement Innovation et Expérimentation (CARDIE), la Délégation académique (ou de région académique) au numérique pour l'éducation (DANE ou DRANE), les corps d'Inspection, les chefs d'établissements scolaires. *In fine*, une dynamique communautaire peut se développer,

englobant la communauté scientifique et la communauté éducative au sens large – une dynamique de recherche qui s'apparente aux concepts anglosaxons de « *community-based research* » (Jason & Glenwick, 2016) et de « *community-based participatory research* » (Campano *et al.*, 2020; Zuber-Skerritt *et al.*, 2020), évoqués par (Pentecouteau, 2020, p. 11) et susceptibles de favoriser la production de connaissances d'autant plus utiles aux différents acteurs concernés qu'elles ont été co-construites par eux tous (Brulin *et al.*, 2023).

Ce qu'il faut retenir

Face aux difficultés que soulève l'opérationnalisation de la différenciation pédagogique, il paraît indispensable de combiner méthodologies de recherche et expertise des praticiens. Dans cette double perspective, le paradigme de la recherche collaborative en sciences de l'éducation semble pouvoir compléter efficacement l'approche par méthodes mixtes. Ce type de recherche-intervention implique de transcender les traditionnels cloisonnements entre visées heuristique, transformative, participative et émancipatrice de la recherche, pour adopter une posture épistémologique propice au métissage des différents types de savoirs portés par le chercheur-intervenant et ses partenaires. Cette approche peut inclure une variété d'acteurs, dont des acteurs institutionnels, pour conférer à la recherche collaborative une dimension communautaire, grâce à laquelle les connaissances produites peuvent s'avérer d'autant plus utiles à l'ensemble des acteurs qu'elles ont été co-construites par eux tous.

PARTIE II : PROBLEMATIQUE ET HYPOTHESES

1 Problématique

Après avoir détaillé le contexte scientifique dans lequel se situe cette thèse, on peut mettre en évidence les éléments de problématisation suivants :

- Relativement à l'accroissement des *inégalités socioscolaires* dans le système éducatif français, le *collège* représente un défi particulier, notamment en raison des *hétérogénéités* qui le caractérisent.
- Dans ce contexte, la *différenciation pédagogique* (DP) est considérée comme une approche porteuse, cependant elle s'avère *difficile à opérationnaliser*, à la fois dans des *études empiriques* et dans les *pratiques de terrain*.
- Les *pédagogies coopératives* constituent un moyen efficace pour mettre concrètement en œuvre la DP au collège, dans la mesure où elles permettent d'accroître l'*autorégulation des apprentissages*, la *métacognition* et l'*auto-efficacité*, *i.e.* le concept intégratif d'*autodirection des apprentissages*.
- A cet égard, les pratiques évaluatives constituent un enjeu de taille ; or la *coévaluation* permet précisément une approche coopérative de l'évaluation, propice à la différenciation pédagogique. Cependant, pour être pratiquée de manière individualisée, elle nécessite le recours au numérique.
- La *coévaluation numérique individualisée* peut être conceptualisée avec profit au moyen de la *théorie instrumentale*, qui permet notamment d'articuler les concepts de *genèse instrumentale*, *artéfact*, *schème d'utilisation*, *instrumentation*, *instrumentalisation* et *catachrèse*.
- Transposé aux usages du numérique, ce cadre théorique permet d'y mettre en évidence un *triple enchâssement des artéfacts didactique, pédagogique et technique*, et d'appréhender la notion de *conflit instrumental*, où l'on distingue *trois degrés de conflictualité* : retard, renoncement partiel et renoncement total.
- *Moodle Atelier* est un artéfact numérique permettant de pratiquer l'inter-évaluation au collège ; toutefois pour pratiquer la coévaluation individualisée il est nécessaire d'y opérer une *catachrèse*.

- En lien avec l'étude de la langue, une telle catachrèse permet de coévaluer dans Moodle Atelier des *écrits produits sur traitement de texte*, ce qui peut de surcroît amener les élèves coévaluateurs à produire des *écrits commentatifs*, ou feedbacks. Bien qu'il s'agisse d'un moyen efficace de mettre en œuvre la différenciation pédagogique au collège, il est difficile de trouver des études empiriques portant sur cette pratique encore émergente qu'est la coévaluation numérique individualisée.
- Pour de telles études empiriques, compte tenu de la complexité de la différenciation pédagogique et des difficultés qu'elle pose en termes d'opérationnalisation pour les chercheurs, la *recherche par méthodes mixtes* (RMM) constitue un choix épistémologique adapté car elle permet de mobiliser de manière intégrative des approches quantitatives et qualitatives.
- En outre, dans la mesure où cet objet complexe qu'est la DP pose également des problèmes de mise en œuvre aux praticiens, il s'agit d'articuler pragmatiquement méthodologie scientifique et pratiques enseignantes en s'appuyant sur le paradigme de la *recherche collaborative* en sciences de l'éducation et de la formation.

A la lumière de ces éléments, nous pouvons formuler la problématique suivante :

« Dans quelles mesure une situation d'enseignement/apprentissage didactisant la production écrite sur traitement de texte au moyen d'un dispositif de coévaluation numérique individualisée permettrait-elle de mettre en œuvre la différenciation pédagogique au collège ? »

2 Variables générales

Cette problématique met en jeu plusieurs *variables générales*, qui déterminent notre hypothèse générale. Dans le contexte du collège, qui est celui de la présente recherche, ces variables générales sont les suivantes :

- Variable d'entrée (*i.e.* variable indépendante) :
 - « Instrumentalisation, au moyen d'une catachrèse, d'un artéfact numérique afin de pratiquer la coévaluation individualisée d'écrits produits sur traitement de texte »
- Variable de sortie (*i.e.* variable dépendante) :

- Différenciation pédagogique

3 Hypothèse générale

En réponse à la problématique énoncée *supra*, l'hypothèse générale qui découle de ces variables générales est la suivante :

« L'instrumentalisation, au moyen d'une catachrèse, d'un artéfact numérique, afin de pratiquer la coévaluation individualisée d'écrits produits sur traitement de texte, permet d'opérationnaliser la différenciation pédagogique. »

4 Variables opérationnelles

Pour opérationnaliser les variables générales que cette hypothèse générale met en jeu, le cadre théorique développé dans notre état de l'art permet de mettre en évidence plusieurs variables opérationnelles (*i.e.* observables/indicateurs).

4.1 Indicateurs en lien avec la théorie instrumentale

Précisons pour commencer que l'« instrumentalisation, au moyen d'une catachrèse, d'un artéfact numérique afin de pratiquer la coévaluation individualisée » est multifactorielle dans la mesure où cette instrumentalisation/catachrèse mobilise le triple enchâssement des artéfacts didactique, pédagogique et technique. Cette dimension multifactorielle sera donc prise en compte en s'appuyant sur la théorie instrumentale, dont découlent les indicateurs suivants :

- Conflits instrumentaux portant sur :
 - L'artéfact didactique ;
 - L'artéfact pédagogique ;
 - L'artéfact numérique ;
 - Une combinatoire de deux ou trois des artéfacts précités.
- Degrés de conflictualité instrumentale :
 - Retards :
 - d'enseignement ;
 - d'apprentissage ;

- Renoncements partiels :
 - d'enseignement ;
 - d'apprentissage ;
- Renoncements totaux :
 - d'enseignement ;
 - d'apprentissage.
- Régulations instrumentales mobilisant :
 - l'artéfact didactique ;
 - l'artéfact pédagogique ;
 - l'artéfact technique ;
 - une combinatoire de deux ou trois des artéfacts précités.

4.2 Indicateur en lien avec la production écrite sur traitement de texte

Par ailleurs, la coévaluation individualisée d'écrits produits sur traitement de texte implique l'indicateur suivant :

- Degré de maîtrise des compétences cibles en maîtrise de la langue et production d'écrit ;

4.3 Indicateurs en lien avec l'autodirection des apprentissages

Ensuite, d'après notre état de l'art, l'opérationnalisation de la différenciation pédagogique sous-entend la mise en œuvre des pédagogies coopératives et de l'autodirection des apprentissages, *i.e.* métacognition et sentiment d'efficacité personnelle et collective. Concrètement, cela implique de mobiliser les indicateurs suivants :

- Degré de maîtrise des métacompétences cibles (*i.e.* métacognition)
- Sentiment d'efficacité personnelle
- Auto-efficacité perçue pour :
 - la réussite scolaire
 - l'autorégulation des apprentissages
 - les apprentissages coopératifs

- Sentiment d'efficacité collective (indice agrégatif personnel)

4.4 Observable en lien avec le paradigme de la recherche collaborative

Enfin, dans la mesure où il s'agit d'une recherche-intervention à caractère collaboratif, dans laquelle nous visons à articuler la posture de chercheur (« recherche sur ») et celle de facilitateur (« intervention pour »), une autre variable indépendante à prendre en compte est le *degré d'accompagnement des enseignants partenaires par le chercheur-intervenant*, relativement à l'instrumentalisation d'un artéfact numérique pour pratiquer la coévaluation individualisée avec leurs élèves.

5 Situation choisie

Pour pouvoir répondre à la problématique énoncée *supra*, et compte tenu des variables opérationnelles afférentes, nous avons déterminé la situation de recherche détaillée ci-après.

Le terrain choisi est le collège public, pour bénéficier d'un accès à l'artéfact numérique Moodle Atelier via l'environnement numérique de travail (ENT) académique. Dans la mesure du possible, on ciblera des collèges de natures variées – *e.g.* urbain, rural, réseau d'éducation prioritaire (REP), péri-urbain –, éventuellement situés dans plusieurs académies ou régions académiques en France.

Pour une recherche collaborative, au-delà du collège stricto sensu, il s'agit de créer des partenariats avec des partenaires institutionnels tels que :

- La Cellule académique recherche développement innovation pour l'éducation (CARDIE)
- Le corps d'Inspection des Lettres
- La Délégation académique (ou de région académique) au numérique éducatif (DANE ou DRANE)

En outre, nous avons vu *supra* qu'en didactique du français les approches croisées de la production d'écrit et de l'étude de la langue constituent une entrée pertinente au collège, spécifiquement au cycle 4. Le public cible – *i.e.* des collégiens de cycle 4 (5^e, 4^e et 3^e) – est donc âgé de 13 ans en moyenne. Outre les considérations scolaires/académiques mentionnées *supra*, signalons que la littérature en psychopédagogie attribue à cette période de la

préadolescence une appétence particulière pour la coopération avec les pairs (Slavin, 1996 ; Yu *et al.*, 2022) – ce qui fait de ce public cible un choix pertinent relativement aux variables opérationnelles suivantes :

- Auto-efficacité perçue pour les apprentissages coopératifs ;
- Sentiment d'efficacité collective.

Enfin, nous avons vu que pour pratiquer la coévaluation individualisée, la situation choisie implique nécessairement l'usage d'un artefact numérique. Dans la perspective d'une articulation efficace entre métacognition et différenciation pédagogique notamment, (Grangeat, 1999) évoque la nécessité pour l'élève de disposer d'un degré d'autonomie suffisant pour choisir entre plusieurs démarches d'apprentissage possibles. Il faut aussi selon lui que ce dernier puisse connaître les effets de ses choix. A ce titre, l'apprenant ne dispose pas ici d'une réelle marge de manœuvre ; à moins que l'enseignant ne scénarise en amont des aménagements (remédiations, approfondissements, possibilité de retravail d'un écrit...) susceptibles de répondre à ses besoins individuels. En revanche, l'abondance de feedback généré par un artefact numérique comme Moodle Atelier (*e.g.* grille de coévaluation critériée, écrit commentatif), à la fois en provenance de l'enseignant et des pairs, est susceptible de permettre à l'élève de connaître les effets de ses choix.

Conformément à notre état de l'art, il s'agit également pour l'élève de parvenir à mobiliser plusieurs modalités de formulation ou de schématisation des informations. A cet égard, la production écrite stricto sensu et la formation d'un écrit commentatif adapté, pertinent et argumenté, à destination des pairs coévalués, constituent deux modes complémentaires de formulation des informations – le deuxième mode consistant en une approche métacognitive, complémentaire du premier. Suivant l'avis de (Perrenoud, 2014), nous avons vu par ailleurs que l'apprenant doit pouvoir « comparer ses propres stratégies cognitives à celles de ses pairs ou s'approprier celles d'un médiateur » (Grangeat, 1999, p. 126). Or un dispositif expérimental de coévaluation numérique fondé sur un artefact numérique comme Moodle Atelier permettrait de favoriser ces interactions coopératives, singulièrement par le truchement de relations d'aide, d'entraide ou de travail de groupe (Connac, 2017b; Freinet, 1980; Reynaud, 2022). En somme, un dispositif expérimental permettant de pratiquer la coévaluation au moyen d'un artefact numérique comme Moodle Atelier correspondrait à une démarche de différenciation pédagogique s'appuyant sur la coopération, mais aussi sur la métacognition, au sens de (Grangeat, 1999 ; Perrenoud, 2014).

6 Prédications

Compte tenu de la situation choisie, les prédictions (*i.e.* les hypothèses opérationnelles ou HO) qui découlent de nos variables sont les suivantes :

6.1 HO 1 : apprentissages en français langue première

« La coévaluation numérique¹⁹ permet à des collégiens de cycle 4 d'améliorer leurs compétences linguistiques pour la production d'écrit. »

6.2 HO 2 : autodirection des apprentissages

« La coévaluation numérique permet à des collégiens de cycle 4 de mieux autodiriger leurs apprentissages. »

Or comme nous l'avons vu *supra* dans notre état de l'art, l'autodirection des apprentissages est un concept intégratif articulant autorégulation, métacognition et auto-efficacité. On peut ainsi décliner HO 2 comme suit :

6.2.1 HO 2a : autorégulation des apprentissages

« La coévaluation numérique permet à des collégiens de cycle 4 de mieux autoréguler leurs apprentissages. »

6.2.2 HO 2b : métacognition

« La coévaluation numérique permet à des collégiens de cycle 4 d'améliorer leur métacognition. »

6.2.3 HO 2c : auto-efficacité perçue

« La coévaluation numérique permet à des collégiens de cycle 4 d'augmenter leur sentiment d'efficacité personnelle. »

¹⁹ En référence à notre état de l'art, il s'agit plus précisément de *pratiquer la coévaluation individualisée au moyen d'un artefact numérique instrumentalisé par catachrèse*. Pour la lisibilité du présent manuscrit, nous avons cependant abrégé la formulation. Cette remarque vaut pour toutes les hypothèses opérationnelles énoncées infra.

6.3 HO 3 : apprentissages coopératifs

« La coévaluation numérique permet de favoriser l'apprentissage coopératif chez des collégiens de cycle 4. »

6.4 HO 4 : conflits et régulations instrumentaux en EIAH

« La coévaluation numérique génère des conflits instrumentaux, mais ces derniers peuvent être régulés »

Or comme nous l'avons vu *supra* dans notre état de l'art, les conflits et régulations instrumentaux en EIAH peuvent concerner la situation d'apprentissage (*i.e.* point de vue de l'apprenant) et/ou la situation d'enseignement (*i.e.* point de vue de l'enseignant). On peut ainsi décliner HO 4 comme suit :

6.4.1 HO 4a : conflits et régulations instrumentaux affectant les situations d'apprentissage

« La coévaluation numérique génère des conflits instrumentaux affectant les situations d'apprentissage, mais ces derniers peuvent être régulés. »

6.4.2 HO 4b : conflits et régulations instrumentaux affectant les situations d'enseignement

« La coévaluation numérique génère des conflits instrumentaux affectant les situations d'enseignement, mais ces derniers peuvent être régulés. »

6.5 Prédications relatives à la fiabilité des instruments

Des prédictions subalternes (codées « PS »), sont prises en compte dans la présente recherche. Elles ont pour fonction de nous renseigner sur la validité des données recueillies, notamment celle des données autorapportées issues des échelles psychométriques – *i.e.* SEPE, SEP(scol), SEP(srl) et SEP(coop).

Ces prédictions sont les suivantes :

- PS 1 : Les élèves participants ont déjà expérimenté l'auto-évaluation et/ou l'inter-évaluation.
- PS 2 : Le questionnaire SEPE est bien accepté par les élèves.

- PS 3 : Les élèves ont bien compris le fonctionnement de l'échelle SEPE.
- PS 4 : La cohérence interne de l'échelle SEPE est satisfaisante.

Les résultats relatifs à ces quatre prédictions constituent un préalable utile à l'analyse de l'ensemble des résultats de cette recherche, cependant ils revêtent un caractère subalterne et sont donc consultables en Annexes 1.1.

PARTIE III : CHOIX METHODOLOGIQUES

Dans cette partie, nous détaillons nos choix méthodologiques, en présentant successivement :

- L'approche épistémologique de la présente recherche ;
- Les participants ;
- Le protocole de recherche ;
- Les outils de recueil et les modalités de traitement des données ;
- Notre « *apparatus* », *i.e.* l'ensemble du matériel informatique utilisé ;
- Les aspects déontologiques qui présidaient à cette étude.

1 Approche épistémologique

Pour tester les hypothèses opérationnelles énoncées *supra*, il s'agit d'envisager la coévaluation numérique individualisée de la production écrite sur traitement de texte au prisme de la théorie instrumentale, avec comme objectif l'étude de la coopération et de l'autodirection des apprentissages en tant que déterminants de la différenciation pédagogique au cycle 4 du collège. Pour ce faire, nous proposons d'employer une méthodologie reposant sur l'utilisation des méthodes mixtes d'une part, et sur le paradigme de la recherche collaborative d'autre part. Notre approche épistémologique est détaillée plus avant dans les paragraphes qui suivent.

1.1 Une expérimentation semi-contrôlée en situation écologique

Pour tester nos hypothèses opérationnelles et répondre à notre question de recherche, nous appliquons les principes de la méthode hypothético-déductive à travers la mise en œuvre d'un protocole semi-expérimental en situation écologique, *i.e.* en classe de français, avec des collégiens de cycle 4 et leurs professeurs de Lettres. Les arguments qui déterminent le choix de cette situation de recherche sont détaillés *supra*.

1.2 Une recherche-intervention collaborative

Comme indiqué *supra*, la présente étude s'inscrit dans le paradigme de la recherche-intervention, et plus précisément dans celui de la recherche collaborative en sciences de l'éducation et de la formation. Il s'agit donc de conduire la recherche en étroite collaboration avec des enseignants *volontaires* et leurs élèves, ainsi qu'avec l'accord et le soutien de partenaires institutionnels tels que la CARDIE, la DANE/DRANE, les corps d'Inspection et les chefs d'établissements. Pour rappel, en réponse aux difficultés d'opérationnalisation qui caractérisent l'objet d'étude lui-même – *i.e.* la différenciation pédagogique –, il s'agit de conférer à la présente recherche une dimension communautaire et participative (« *community-based participatory research* »). Les apports et les implications de cette approche épistémologique, tels qu'ils se sont concrètement actualisés dans le vécu de la recherche, font l'objet d'un chapitre dédié, *infra*, dans la Discussion générale (partie V).

1.3 Une recherche par méthodes mixtes (RMM)

Suivant les préconisations, détaillées *supra*, de (Forget, 2017b, p. 50), la mise en œuvre de la RMM dans la présente étude repose précisément sur les cinq thèmes précédemment cités, qui émergent de l'analyse par comparaison réalisée par (Johnson *et al.*, 2007), *i.e.* :

1. Les types de données susceptibles d'être mélangées ;
2. L'étape de la recherche où le mélange a lieu ;
3. L'ampleur de la RMM
4. Les raisons qui orientent le mélange ;
5. L'orientation de la RMM.

Pour commencer, concernant les données numériques et textuelles recueillies, nous complétons nos analyses quantitatives avec des analyses qualitatives (thèmes 1 et 5), sachant que notre phase d'intégration des résultats intervient en aval de la recherche (thème 2), par *combinaison* (Anadón, 2019, pp. 110-111). En d'autres termes, les données textuelles et numériques sont collectées simultanément selon les deux approches, puis analysées séparément, avant de fusionner les résultats (Figure 16).

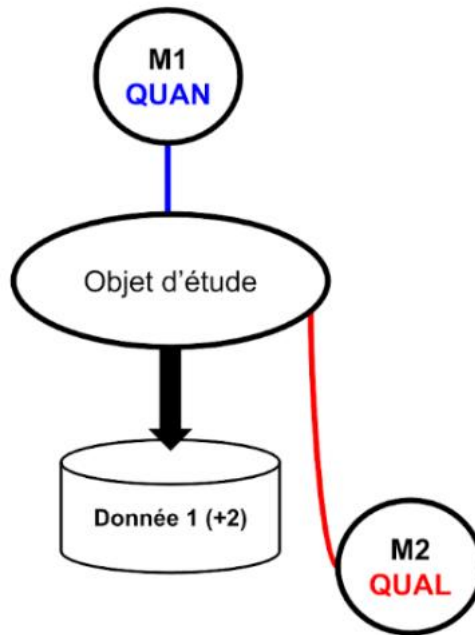


Figure 16 – Modèle d'intégration par combinaison

La comparaison qui en résulte a vocation à renforcer le processus de validation des hypothèses opérationnelles, en croisant les résultats issus des approches quantitatives et qualitatives. Précisons que dans la présente recherche les deux approches n'ont pas tout à fait le même statut : l'approche qualitative joue un rôle important mais subsidiaire, relativement à l'approche quantitative, qui est la méthode principale (thèmes 3 et 4). Le modèle (ou devis) adopté pour l'intégration des résultats quantitatifs et qualitatifs est le modèle parallèle et convergent tel que (Creswell & Creswell, 2017, p. 214) l'ont conceptualisé (Figure 17).

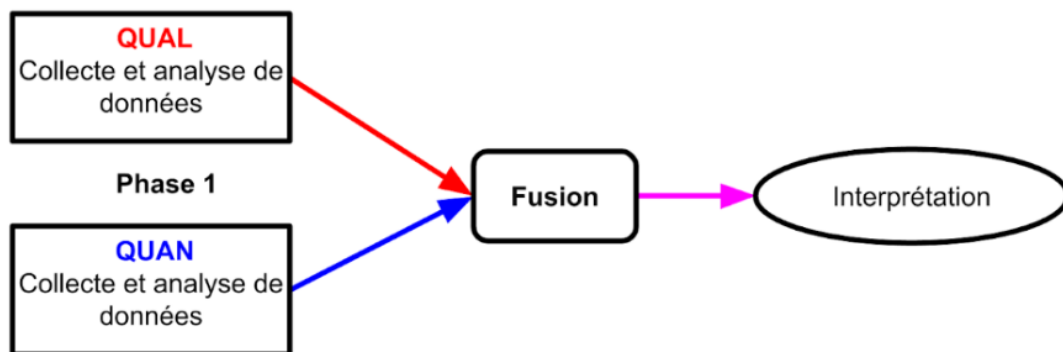


Figure 17 – Modèle parallèle et convergent d'après (Creswell & Creswell, 2017, p. 210)

2 Participants

Cette recherche-intervention de type collaboratif implique différents types d'acteurs et de partenaires institutionnels, que nous présentons ici en détail.

2.1 Collèges

Cette recherche-intervention se déroule au collège, plus précisément dans trois collèges publics (EPLE²⁰) de natures variées :

- Le collège A, situé en zone rurale, dans le Gard (30) ;
- Le collège B situé en zone urbaine (Paris), en réseau d'éducation prioritaire (REP) ;
- Le collège C situé en zone péri-urbaine, près de Montpellier (34).

2.2 Académies et partenaires institutionnels afférents

Deux académies participent à la présente recherche :

- Académie de Montpellier
- Académie de Paris

Dans ces deux académies, les partenaires institutionnels suivants collaborent au projet :

- La Cellule académique recherche développement innovation pour l'éducation (CARDIE)
- L'Inspection des Lettres
- La Délégation académique au numérique éducatif (DANE) de Paris
- La Délégation de région académique au numérique éducatif (DRANE) d'Occitanie

2.3 Enseignantes partenaires

Dans chacun des trois collèges partenaires, 1 professeure de Lettres modernes se porte volontaire pour devenir partenaire de cette recherche collaborative, avec à chaque fois une ou deux classe(s) de cycle 4, pour un total de 109 participants²¹, âgés de treize ans en moyenne :

- Enseignante A (collège A) :
 - Classe A (3^e), 14 élèves ;

²⁰ EPLE : Etablissement public local d'enseignement.

²¹ 109 élèves ont participé à l'étude, mais seulement 80 dans le cadre du protocole détaillé infra. Les 29 autres élèves ont expérimenté le dispositif hors protocole ; cependant cela a pu être partiellement documenté dans la présente recherche.

- Classe A' (4^e), 19 élèves ;
- Enseignante B (collège B) :
 - Classe B (3^e), 24 élèves ;
 - Classe B' (3^e)²², 29 élèves ;
- Enseignante C (collège C) :
 - Classe C (5^e), 23 élèves ;
 - Classe C' (6^e), 24 élèves²³ (

L'entièreté du cycle 4 (5^e, 4^e, 3^e) est donc représentée dans la présente recherche.

Par ailleurs, les trois enseignantes partenaires sont des enseignantes très expérimentées, ayant exercé pendant au moins 20 ans. Toutes sont impliquées dans la formation initiale et continue des enseignants et ont l'habitude de travailler en équipe, notamment avec le corps d'Inspection des Lettres.

2.4 Échantillons

Pour les besoins de l'analyse statistique notamment, chaque classe ayant expérimenté le dispositif de coévaluation numérique (y compris hors protocole) est considérée comme un échantillon :

- Echantillon A (*i.e.* classe A, niveau 3^e, 14 élèves) ;
- Echantillon A' (*i.e.* classe A', niveau 4^e, 19 élèves) ;
- Echantillon B (*i.e.* classe B, niveau 3^e, 24 élèves) ;
- Echantillon B' (*i.e.* classe B', niveau 3^e, 29 élèves) ;
- Echantillon C (*i.e.* classe C, niveau 5^e, 23 élèves).

Par ailleurs, un « super-échantillon » (SE) permet de regrouper, le cas échéant, l'ensemble des échantillons ayant participé au protocole semi-expérimental, *i.e.* les échantillons A, A', B et C. La justification de ce choix, relatif au traitement statistique des données, est fournie *infra*.

²² Il s'agit des 29 élèves précités, qui ont expérimenté la coévaluation numérique en année N+1, hors protocole.

²³ Cette classe ne participe pas au protocole semi-expérimental, mais seulement à un focus groupe dont les détails sont indiqués *infra*.

3 Protocole de recherche

Le protocole de la présente recherche est centré sur la mise en œuvre semi-expérimentale, en situation écologique, d'une activité de coévaluation instrumentée par le numérique. Les principales caractéristiques de ce protocole sont résumées dans la Figure 18 et dans la section qui suit. Nous examinerons ensuite plus en détail ses composantes didactique, pédagogique et numérique, avant de préciser la durée de la période expérimentale et le nombre d'itérations du protocole pendant cette période.

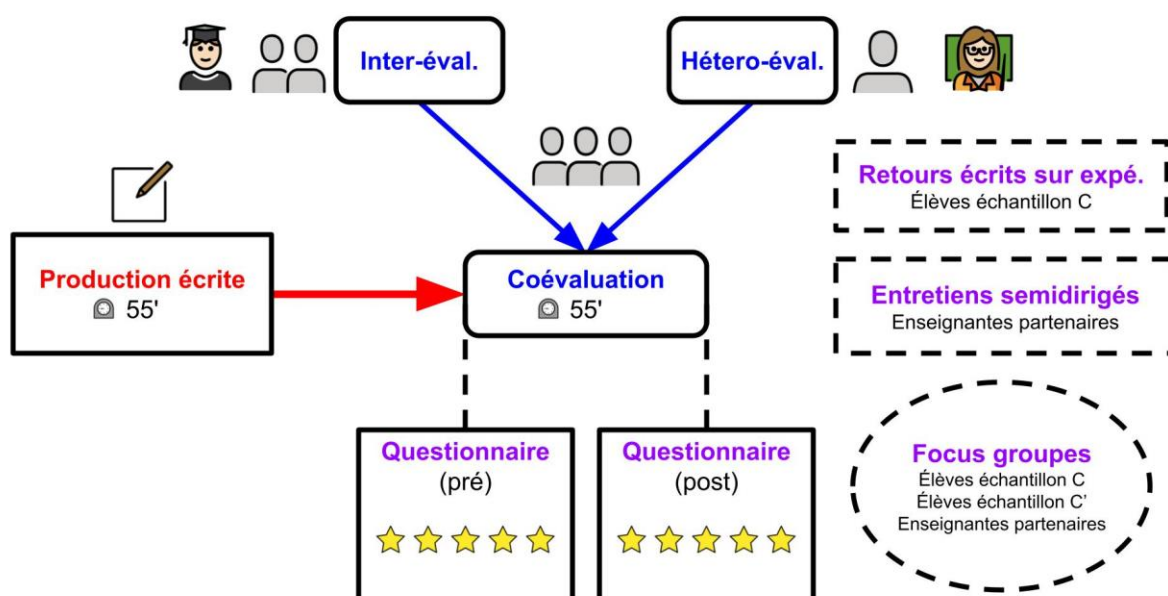


Figure 18 – Représentation schématique du protocole expérimental

3.1 Principales caractéristiques

- Le protocole de recherche est partiellement co-construit avec les enseignantes partenaires, notamment pour ce qui relève de la didactisation de la coévaluation formative individualisée en EIAH.
- Cette didactisation de la coévaluation formative individualisée, instrumentée par Moodle Atelier, constitue l'élément central du protocole.

- Lors d'une séance de 55 minutes en salle informatique, elle a pour objet la coévaluation d'écrits courts, que les élèves ont produit en classe en amont, sur traitement de texte²⁴, lors d'une séance préalable durant également 55 minutes.
- Cette activité de coévaluation génère deux types de données :
 - d'une part un score de scripteur (*i.e.* l'hétéro-évaluation de l'enseignante partenaire) mesurant la maîtrise, par l'élève participant, des compétences cibles en français ;
 - d'autre part un score de coévaluateur mesurant la maîtrise des métacompétences afférentes, *i.e.* la capacité à apprécier avec justesse les compétences du pair coévalué.
- Dans Moodle Atelier, selon la catachrèse décrite *supra*, c'est l'hétéro-évaluation de l'enseignante partenaire qui sert de point de référence pour mesurer, compétence par compétence, la pertinence des coévaluations d'élèves.
- Selon une approche pré/post, chaque séance de coévaluation numérique est précédée et suivie de la passation d'une échelle psychométrique mesurant le sentiment d'efficacité personnelle des élèves participants (échelle SEPE²⁵).
- Cette échelle SEPE est intégrée à un questionnaire en ligne contenant d'autres items, indépendants, spécifiquement dédiés :
 - à la mesure du sentiment d'efficacité collective (indice agrégatif personnel) ;
 - à la mesure du degré de conflictualité perçue ;
 - au recueil, de façon ouverte (semidirigée), des données textuelles relatives à la genèse instrumentale du dispositif de coévaluation numérique, du point de vue des élèves.
- En outre, avec l'échantillon C seulement²⁶ :
 - un retour écrit sur expérience est recueilli ;
 - deux focus groupes sont organisés.

²⁴ Ces éléments didactiques du protocole sont détaillés infra.

²⁵ SEPE : Sentiment d'efficacité personnelle des enfants, pré-adolescents et adolescents.

²⁶ L'échantillon C est le seul avec lequel nous réitérons le protocole de recherche trois fois ; en outre, le chef d'établissement concerné a consenti à un aménagement de l'emploi du temps, de sorte que les retours écrits sur expérience ont pu être recueillis et les focus groupes mis en œuvre.

- Entre T4 et T5, un focus groupe est également mis en œuvre avec les élèves de la 6^e coopérative du collège C, une autre classe de l’enseignante C, qui ne prenait pas part au protocole semi-expérimental.
- Des entretiens semidirigés et un focus groupe sont par ailleurs mis en œuvre avec les enseignantes partenaires.

3.2 Artéfact didactique : production écrite

La présente recherche donne lieu à onze productions écrites coévaluées, dont neuf en année N, dans le cadre du protocole expérimental détaillé ci-après, et deux hors protocole, en année N+1. Les neuf séances qui se déroulent dans le cadre du protocole sont codées chronologiquement, en lien avec les temps de mesure pré/post afférents, notamment eu égard à la collecte des données psychométriques :

- T1/T2 : 1^{ère} séance expérimentale ;
- T3/T4 : 2^e séance expérimentale ;
- T5/T6 : 3^e séance expérimentale (valable uniquement pour l’échantillon C).

La production écrite s’effectue à chaque fois en salle informatique, sur ordinateur et traitement de texte – sauf dans le cas de l’échantillon C, où une première phase de travail au brouillon est systématiquement mise en œuvre pendant 55 minutes, en amont, de façon manuscrite.

Les onze scénarios de production écrite sont recensés ci-après, en précisant pour chacun la durée, la consigne d’écriture, ainsi que les critères d’évaluation afférents, présentés sous forme de grille (captures d’écrans issues de Moodle Atelier). Selon l’approche par compétences, chaque critère de réussite constitue en quelque sorte une micro-compétence.

A l’exception des séances hors protocole de l’année N+1, la coévaluation se fait selon une échelle de Likert à quatre points, usuelle en contexte scolaire, *e.g.*, par ordre croissant de maîtrise : insuffisant ; peu satisfaisant ; satisfaisant ; très satisfaisant (Figure 19).

Niveaux			
<input type="radio"/> Insatisfaisant	<input checked="" type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant

Figure 19 – Exemple d’échelle de Likert à quatre points pour la coévaluation dans Moodle Atelier

3.2.1 T1/T2

3.2.1.1 Classe A (3^e)

Au début d'une séquence d'enseignement ayant pour objet l'étude intégrale du roman de Didier Daeninckx intitulé *Cannibale* (Folio, 2000), la première itération du protocole expérimental porte sur un extrait de l'œuvre, travaillé à la manière de l'épreuve de français du Diplôme national du brevet (DNB), *i.e.* de « Nous avons embarqué... » (p. 19) à « Hommes anthropophages de Nouvelle-Calédonie » (p. 22). Lors d'une première séance de 55 minutes, en rédigeant un court paragraphe argumenté, les élèves doivent répondre à la question ci-dessous, proposée par l'enseignante A :

« Le narrateur correspond-il à cette image que veut donner la direction ? Rédigez une réponse développée et justifiée, en prenant appui sur l'ensemble du texte et des réponses apportées aux questions précédentes. »

Les critères de réussite – *i.e.* les micro-compétences –, conjointement évalués par l'enseignante A et les élèves coévaluateurs lors d'une seconde séance de 55 minutes, sont les suivants (Figure 20) :

Critères
Repérer et déduire les informations du texte.
S'appuyer sur nos connaissances.
Veiller à faire une réponse complète.
La réponse contient plusieurs idées, organisées grâce à des connecteurs.
Pour chaque élément de réponse, on formule l'idée avec nos mots et on justifie en citant le texte.
On fait des phrases complexes.

Figure 20 – Critères de coévaluation pour la séance T1/T2 avec la classe A (3^e)

En moyenne, chaque élève a 2 écrits de pairs à coévaluer.

3.2.1.2 Classe A' (4^e)

Pour la classe A' (4^e) à T1/T2, la première itération du protocole se situe au début d'une séquence portant sur le théâtre. Lors d'une première séance de 55 minutes, il s'agit d'écrire une critique de la mise en scène par Denis Podalydès (2006) de la pièce de théâtre *Cyrano de Bergerac* d'Edmond Rostand. Cette mise en scène a été préalablement étudiée en classe. Les critères d'évaluation (Figure 21) sont par ailleurs co-construits en amont avec les élèves, selon une approche socio-constructiviste (Annexes 2.1.1). En moyenne, lors d'une seconde séance de 55 minutes, chaque élève a trois critiques de pairs à coévaluer.

Critères
- le texte contient des informations : date de création, acteur, décor, thème, intrigue, metteur en scène, dramaturge... > décrire ce qu'on voit > employer le vocabulaire spécifique au théâtre
- formuler un jugement sur ce qu'on voit (adjectif, verbe d'opinion...)
- forme du texte : paragraphe, texte organisé, connecteurs...
- la critique doit montrer qu'on a bien compris le texte.

Figure 21 – Critères de coévaluation pour la séance T1/T2 avec la classe A' (4^e)

3.2.1.3 Classe B (3^e)

A T1/T2, dans le cadre d'une séquence portant sur la 1^{ère} Guerre mondiale, il s'agit pour les élèves de la classe B (3^e) de rédiger, lors d'une première séance de 55 minutes, un compte-rendu portant sur un exposé d'élève, au choix, parmi ceux qui leur ont été présentés en classe à la séance précédente. En moyenne, lors d'une seconde séance de 55 minutes, chaque élève a trois coévaluations et 1 auto-évaluation à effectuer. Les critères retenus par l'enseignante pour la coévaluation sont les suivants (Figure 22) :

Critères
Respect de la consigne
Reprise des informations et contenus
Organisation des idées
Vocabulaire adapté et varié
Respect des accords S-V
Emploi des pronoms
Utilisation des temps de l'indicatif : conjugaison et emploi
Respect des accords dans le GN
Variété et correction syntaxique des types et formes de phrases
Respect de l'orthographe lexicale
Emploi d'adverbes

Figure 22 – Critères de coévaluation à T1/T2 pour la classe B (3e)

3.2.1.4 Classe C (5^e)

A T1/T2, dans le cadre d'une séquence sur les héros et héroïnes dans l'occident médiéval, les élèves de la classe C (5^e) doivent réaliser, lors d'une première séance de 55 minutes, un écrit projectif, *i.e.* une « greffe » textuelle. Le dernier paragraphe d'un récit de chevalerie étudié en classe leur est rappelé en amont de la consigne (Annexes 2.1.2). En moyenne, lors d'une seconde séance de 55 minutes, chaque élève doit coévaluer trois écrits de pairs. Parmi ces trois travaux, ils peuvent avoir à coévaluer l'écrit de l'enseignante partenaire, car elle s'est « prêtée au jeu » de la coévaluation, au moyen d'un écrit délibérément imparfait, auto-évalué en amont, et sous le couvert de l'anonymat fourni par Moodle Atelier. Les critères de coévaluation sont les suivants (Figure 23) :

Critères
Le lecteur comprend les intentions de la jeune femme, et ce qu'elle ressent.
Des adjectifs mélioratifs ont été utilisés pour mettre la beauté de la femme en valeur. D'autres procédés accentuent cette beauté.
Le bouquet est précisément décrit et a un rôle dans l'histoire.
Le sens de l'odorat a été utilisé, sans exclure la mise en valeur des autres sens.
Le lecteur perçoit les mouvements et les sensations liés au vent.
Le texte contient au moins une comparaison correctement construite (forme) et appropriée (sens).

Figure 23 – Critères de coévaluation à T1/T2 pour la classe C (5e)

3.2.2 T3/T4

3.2.2.1 Classe A (3^e)

A T3/T4 (2^e itération du protocole), dans le cadre d'une séquence portant sur l'objet d'étude « Progrès et rêves scientifiques », les élèves de la classe A (3^e) sont invités lors d'une première séance de 55 minutes à rédiger la définition d'une machine futuriste de leur invention. Leur écrit d'invention doit respecter les critères de coévaluation suivants, co-construits en classe en amont (Figure 24) :

Critères
<p>L'explication semble crédible :</p> <ul style="list-style-type: none"> - le fonctionnement a l'air possible - on s'appuie sur des éléments du réel (éléments scientifiques, vocabulaire précis...)
<p>C'est une machine inventée :</p> <ul style="list-style-type: none"> - le fonctionnement reste imaginaire - quelques néologismes ont été utilisés
<p>Les relations logiques sont bien employées :</p> <ul style="list-style-type: none"> - au moins 6 phrases correctes expriment chacune des 6 relations logiques

Figure 24 – Critères de coévaluation à T3/T4 pour la classe A (3e)

En outre, deux exemples de travaux, créés en amont par l'enseignante, sont fournis dans Moodle Atelier, à l'aide de la fonctionnalité « Travaux exemplaires ». En moyenne, lors d'une seconde séance de 55 minutes, en plus de la coévaluation de ces exemples de travaux, chaque élève doit coévaluer la production écrite de 2 pairs.

3.2.2.2 Classe A' (4^e)

Lors de la séance T3/T4 (2^e itération du protocole), dans le cadre d'une séquence sur le genre fantastique, les élèves de la classe A' (4^e) doivent répondre à deux questions de compréhension type DNB, portant sur la nouvelle *Le Portrait ovale* d'Edgar Poe :

- Quels éléments décrivent le caractère et l'attitude du peintre ? Vous rédigerez une réponse organisée et justifiée, en formulant trois éléments de réponse.
- Le texte suggère deux manières d'expliquer la mort de la jeune femme. Formulez et justifiez ces deux explications.

Dans le cadre d'une première séance de 55 minutes, leur écrit doit respecter les critères d'auto- et de coévaluation suivants, co-construits en classe en amont (Figure 25 et Figure 26) :

Critères
Rédiger une réponse justifiée, qui respecte les 3 étapes : formuler l'idée, l'expliquer, la justifier.
Bien séparer les idées et employer des connecteurs logiques.
Repérer les 3 éléments de réponse attendus et relever des justifications pertinentes : - c'est un bon peintre, un artiste renommé (l. 20, 24-25) - le peintre est passionné par son art, au point qu'il en devient fou (l. 7, 16-17, 27) - il semble devenir négligent, voire égoïste, ne se préoccupant plus de son épouse (l. 17-19, 27-28)

Figure 25 – Critères de coévaluation à T3/T4 pour la classe A' (4e) : 1^{ère} question

Critères
Rédiger une réponse justifiée, qui respecte les 3 étapes : formuler l'idée, l'expliquer, la justifier.
Bien séparer les idées et employer des connecteurs logiques.
Repérer les 2 éléments de réponse attendus et relever des justifications pertinentes : - La femme est morte car elle est négligée par son mari, elle s'affaiblit, elle se languit (l. 17-22) - La femme est morte car le tableau lui a pris sa vie. (l. 27-28, 35-36)

Figure 26 – Critères de coévaluation à T3/T4 pour la classe A' (4e) : 2^{nde} question

Lors d'une seconde séance de 55 minutes, chaque élève doit s'auto-évaluer puis coévaluer la production écrite d'un pair pour chaque question – soit 4 auto- et coévaluations au total.

3.2.2.3 Classe B (3^e)

A T3/T4, dans le cadre d'une séquence portant sur la poésie, il s'agit pour les élèves de la classe B (3^e) de rédiger en 55 minutes une réponse argumentée à une question de compréhension type DNB, portant sur le recueil poétique *Toute la Lyre* de Victor Hugo. En moyenne, lors des 55 minutes de la séance suivante, chaque élève a trois coévaluations et 1 auto-évaluation à effectuer. Les critères retenus par l'enseignante pour l'auto- et la coévaluation sont les suivants (Figure 27) :

Critères
Dans le développement de la réponse, les 3 idées attendues sont présentes : <ul style="list-style-type: none">- conséquences des rumeurs- secrets qui nous échappent- Hugo conseille la prudence et la discrétion
Justification : références au poème
Qualité de l'expression : phrases bien construites
Vocabulaire adapté
Utilisation des temps de l'indicatif : conjugaison et valeurs
Quantité attendue : 5 à 10 lignes

Figure 27 – Critères de coévaluation à T3/T4 pour la classe B (3^e)

3.2.2.4 Classe C (5^e)

Dans le cadre de la même séquence que précédemment, portant sur les héros et héroïnes dans l'occident médiéval, les élèves de la classe C (5^e) doivent réaliser à T3/T4 un écrit d'invention à partir d'une image (Annexes 2.2, Figure 101), en l'occurrence l'œuvre *Windflowers* de John William Waterhouse (1903). La production d'écrit se déroule en 55 minutes et la consigne donnée par l'enseignante C est la suivante :

- Dans un texte de 10 à 15 lignes, strictement,
- vous raconterez où va ce personnage féminin, et ce qu'il ressent,
- en restant sur un temps limité.
- Sa beauté sera mise en valeur.
- Le bouquet d'anémones sera précisément décrit et jouera un rôle dans votre bref récit.

En moyenne, lors de la séance qui suit (55 minutes également), chaque élève doit coévaluer trois écrits de pairs. Comme précédemment, l’enseignante partenaire participe elle aussi en tant que scripteur, au moyen d’un écrit délibérément « imparfait » qu’elle a auto-évalué au préalable. Certains élèves coévaluent donc la production écrite de l’enseignante, mais sous le couvert de l’anonymat fourni par Moodle Atelier. Les critères de coévaluation sont les suivants (Figure 28) :

Critères
Le lecteur comprend les intentions de la jeune femme, et ce qu'elle ressent.
Des adjectifs mélioratifs ont été utilisés pour mettre la beauté de la femme en valeur. D'autres procédés accentuent cette beauté.
Le bouquet est précisément décrit et a un rôle dans l'histoire.
Le sens de l'odorat a été utilisé, sans exclure la mise en valeur des autres sens.
Le lecteur perçoit les mouvements et les sensations liés au vent.
Le texte contient au moins une comparaison correctement construite (forme) et appropriée (sens).

Figure 28 – Critères de coévaluation à T3/T4 pour la classe C (5e)

3.2.3 T5/T6

3.2.3.1 Classe C (5^e)

Pour les élèves de la classe C (5^e), dans le cadre d’une séquence ayant pour objet d’étude « Vivre en société, participer à la société », la troisième itération du protocole expérimental porte sur le roman d’Eric-Emmanuel Schmitt intitulé *Le Sumo qui ne pouvait pas grossir* (Albin Michel, 2009). En 55 minutes, il s’agit pour les élèves de rédiger un récit contenant un dialogue et se greffant à la suite d’un extrait de l’œuvre étudiée en classe (Annexes 2.3, Figure 102). Pour les y aider, une image est fournie en support – *i.e.* la 1^{ère} de couverture du livre (Annexes 2.3, Figure 103). La consigne d’écriture proposée par l’enseignante C est la suivante :

« Lors de son arrivée à l'école de Shomintsu, Jun a l'occasion de parler avec le champion, Ashoryu. Celui-ci l'accueille et lui donne des conseils pour bien s'intégrer et pour affronter la rude préparation des sumos.

En une vingtaine de lignes (entre 18 et 22 lignes), vous rendrez compte de cette discussion, en utilisant le discours direct, bien disposé et bien ponctué.

L'ajout d'un autre personnage est autorisé, mais pas davantage. La description d'un des personnages est requise.

Pour vous aider à la description d'un personnage, vous pouvez vous inspirer de la première de couverture de l'œuvre, ci-après²⁷. »

En moyenne, les élèves ont trois coévaluations à effectuer lors de la séance suivante (55 minutes). Là encore, certains d'entre eux se voient attribuer l'écrit délibérément imparfait que l'enseignante partenaire a écrit et auto-évalué dans Moodle Atelier en amont, sous le couvert de l'anonymat fourni par le logiciel. La grille de coévaluation, co-construite en amont par l'enseignante C et ses élèves, est la suivante (Figure 29) :

Critères
Le récit contient un dialogue au discours direct, correctement ponctué et disposé.
La narration ou le dialogue contient au moins deux verbes conjugués au conditionnel.
Le texte contient au moins deux figures d'analogie de nature différente : comparaison (figure de style et / ou degré de l'adjectif) ; métaphore.
Le texte comportera des indications sur le physique et sur le caractère d'un des personnages au moins.
Des éléments de la culture japonaise seront perceptibles dans le texte.

Figure 29 – Critères de coévaluation à T5/T6 pour la classe C (5e)

3.2.4 Hors protocole (année N+1)

A deux reprises, lors de l'année N+1 (2022-2023), avec une nouvelle classe B' (3^e) – distincte de l'échantillon B avec lequel expérimentons lors de l'année N –, l'enseignante B pratique la coévaluation numérique individualisée avec Moodle Atelier, en semi-autonomie – *i.e.* sans intervention du chercheur en amont de la mise en œuvre, mais avec une intervention marginale par la suite – pour réguler certains conflits instrumentaux. En lien avec notre hypothèse opérationnelle HO 4, portant sur les conflits instrumentaux et les régulations afférentes, ces 2 initiatives font l'objet d'une analyse détaillée *infra*, dans une section ad hoc de notre partie « Résultats » (partie IV). Les artéfacts didactiques sont quant à eux détaillés ici, dans les deux sections qui suivent.

²⁷ La première de couverture a été téléversée dans Moodle par l'enseignante C et les élèves pouvaient y accéder d'un simple clic.

3.2.4.1 Classe B' (3^e) : 1^{ère} itération semi-autonome

Concernant la première activité ainsi mise en œuvre, s'agissant d'une classe de 3^e et dans la perspective de l'épreuve écrite de français du diplôme national du brevet (DNB), l'objectif didactique de l'enseignante B est d'articuler maîtrise de la langue et production écrite, en l'occurrence en mettant en œuvre une activité de production écrite brève, en lien avec l'actualité, singulièrement pour exercer la vigilance des élèves à l'égard des critères d'évaluation afférents.

Dans le cadre d'une séquence portant sur l'autobiographie, le choix de l'enseignante se porte sur un support bref, *i.e.* un extrait de captation vidéo d'une interview radiophonique d'Annie Ernaux – prix Nobel de littérature 2022 – sur France Culture, d'une durée inférieure à 5 minutes. A travers ce choix de support, l'enseignante B souhaite mettre l'accent sur la potentielle attractivité du support vidéo, combinée à celle de la coévaluation et de l'auto-évaluation (artéfacts pédagogiques), par l'entremise de l'outil informatique – *i.e.* Moodle Atelier (artéfact technique) –, pour favoriser chez les élèves :

- une écriture désinhibée, en vue d'un focus accru sur les caractéristiques formelles de leur production écrite (*i.e.* réponse courte à une question de compréhension type DNB) ;
- une vigilance métacognitive à l'égard des critères d'évaluation afférents ;
- une vigilance vis-à-vis de la correction orthographique et morphosyntaxique de la langue.

Les trois composantes précitées constituent, *in fine*, l'artéfact didactique. En outre, compte tenu du fait que, globalement, les compétences de cette classe en compréhension orale sont plutôt fragiles, et bien que le registre langagier de cette interview soit relativement accessible, l'enseignante B a mis en place, dans son cours Moodle – celui-là même où se trouve également l'Atelier Moodle –, un questionnaire à choix multiple (QCM) autocorrectif. Cette précaution vise à garantir que l'accès à l'artéfact didactique décrit *supra* n'est pas gêné par des difficultés de compréhension liées au support utilisé.

Dans le cadre d'une première séance de 55 minutes, la consigne d'écriture est la suivante :

« Rédigez un texte de 10 à 15 lignes qui résume ce que vous avez retenu des propos d'Annie Ernaux. Aidez-vous des questions et des réponses au quiz. »

En moyenne, les élèves doivent effectuer 4 coévaluations, dont 1 auto-évaluation. La coévaluation se fait au moyen d'une échelle de Likert, mais il s'agit ici de descriptifs ad hoc, formulés spécifiquement pour chaque critère. La grille de coévaluation proposée par l'enseignante B est présentée ci-dessous (Figure 30) :

Critères	Niveaux			
Reprise des informations données dans l'entretien.	<input type="radio"/> Aucune information.	<input type="radio"/> Il manque de nombreuses informations.	<input type="radio"/> Il y a la plupart des informations.	<input type="radio"/> Toutes les informations sont reprises.
Respect du nombre de lignes attendues (10 à 15)	<input type="radio"/> Il y a moins de 10 lignes.		<input type="radio"/> Le texte comporte de 10 à 15 lignes.	
Le vocabulaire est bien adapté et varié.	<input type="radio"/> Le vocabulaire est pauvre et n'est pas adapté.		<input type="radio"/> Le vocabulaire est varié et bien adapté	
Les phrases sont brèves et bien construites.	<input type="radio"/> Les phrases sont mal construites et ne respectent pas la syntaxe.	<input type="radio"/> Il y a quelques maladresses de syntaxe et de ponctuation	<input type="radio"/> Les phrases respectent la syntaxe	
Les conjugaisons sont maîtrisées et les temps bien employés.	<input type="radio"/> Nombreuses erreurs de conjugaisons et d'emplois des temps	<input type="radio"/> Quelques erreurs	<input type="radio"/> Les conjugaisons sont maîtrisées et les temps bien employés.	
Les accords sont respectés (sujets verbes / adjectifs et noms / participes....)	<input type="radio"/> Nombreuses erreurs	<input type="radio"/> Quelques erreurs	<input type="radio"/> Pas d'erreurs	

Figure 30 – Grille de coévaluation pour la classe B' (3^e, année N+1, 1^{ère} itération hors protocole)

3.2.4.2 Classe B' (3^e) : 2^{nde} itération semi-autonome

La seconde réappropriation autonome du dispositif par l'enseignante B avec l'échantillon B' lors de l'année N+1 se situe dans le cadre d'une séquence sur la Première Guerre mondiale, à partir de 4 bandes-annonces de films en lien avec cette thématique, que les élèves sont invités à visionner chez eux en amont. En outre, pour chaque film, les élèves doivent accomplir à la maison un travail préparatoire, qui consiste à remplir une fiche signalétique (*e.g.* titre du film, nom et intentions du réalisateur). Selon un scénario pédagogique hybride, articulant distanciel et présentiel, l'objectif est de les entraîner à la rédaction d'un brouillon ou « écrit de travail », puis de s'appuyer sur cet écrit préparatoire pour produire un paragraphe rédigé, en présentiel, en salle informatique, lors d'une première séance de 55 minutes. La coévaluation de ces écrits dans Moodle Atelier intervient ensuite, lors d'une seconde séance de durée égale.

La consigne d'écriture est la suivante :

« Vous allez rédiger un paragraphe de dix lignes pour présenter la bande-annonce de film que vous avez choisie. »

Il faut reprendre les informations de la fiche que vous avez préparée. »

En moyenne, les élèves doivent effectuer 6 coévaluations, dont 1 auto-évaluation. Là encore, l'auto-évaluation et la coévaluation se font au moyen d'une seule et même échelle de Likert, mais avec des descriptifs ad hoc, formulés spécifiquement pour chaque critère. La grille d'auto- et de coévaluation proposée par l'enseignante B est présentée ci-dessous (Figure 31) :

Critères	Niveaux		
Le paragraphe comporte dix lignes (police Arial ou Colibri 11).	<input type="radio"/> 0 à 4 lignes	<input type="radio"/> 4 à 8 lignes	<input type="radio"/> 9 à 11 lignes
Les informations de la fiche sont toutes présentes.	<input type="radio"/> Aucune information	<input type="radio"/> Quelques informations sont reprises mais il en manque.	<input type="radio"/> Toutes les informations sont reprises.
Le vocabulaire est varié et adapté.	<input type="radio"/> Le vocabulaire est pauvre, répétitif ou d'un niveau de langue inadapté.	<input type="radio"/> Le vocabulaire est correct mais peu varié.	<input type="radio"/> Le vocabulaire est varié, riche et très bien adapté au propos.
Les phrases sont bien construites et respectent la syntaxe.	<input type="radio"/> De nombreuses erreurs de construction de phrases.	<input type="radio"/> Quelques maladresses de constructions.	<input type="radio"/> Des phrases bien construites. La syntaxe est respectée.
Les accords sujets-verbes sont respectés.	<input type="radio"/> Plus de 4 erreurs	<input type="radio"/> de 1 à 4 erreur(s).	<input type="radio"/> Aucune erreur
Les accords de participes passés et d'adjectifs sont respectés.	<input type="radio"/> Plus de 5 erreurs	<input type="radio"/> Entre 1 et 5 erreur(s).	<input type="radio"/> Aucune erreur.
L'orthographe lexicale est respectée.	<input type="radio"/> Plus de 5 erreurs.	<input type="radio"/> Entre 1 et 5 erreur(s).	<input type="radio"/> Aucune erreur.

Figure 31 – Grille de coévaluation pour la classe B' (3^e, année N+1, 2nde itération hors protocole)

3.3 Artéfact pédagogique : coévaluation

Pour mettre en œuvre la coévaluation individualisée, notre étude s'inspire de notre état de l'art pour proposer une activité de coévaluation, *i.e.* un agrégat d'hétéro-évaluation et d'inter-évaluation, avec une part importante de dialogue et de négociation entre apprenants et enseignants, en amont et en aval de l'activité de coévaluation proprement dite.

Il est à noter que l'auto-évaluation joue seulement un rôle marginal dans notre dispositif : elle revêt un caractère optionnel et n'est que rarement mobilisée²⁸ par les enseignantes participantes, tandis que la coévaluation est, quant à elle, au cœur du processus.

²⁸ Dans le cadre de ce protocole co-construit, l'auto-évaluation n'est mobilisée que par les enseignantes partenaires B à T2 et T4 avec l'échantillon B, et A à T4 avec l'échantillon A'.

3.4 Artéfact numérique : Moodle Atelier

Dans notre recherche, l'artéfact technique/numérique Moodle Atelier constitue un outil de recueil de données à part entière dans la mesure où il permet de collecter de façon centralisée plusieurs types de données textuelles et numériques :

- Les productions écrites des élèves ;
- Les scores de scripteurs (issus de l'hétéro-évaluation de l'enseignant) ;
- Les scores de coévaluateurs ;
- Les consignes des enseignants ;
- Les éléments de paramétrage (instrumentalisation et catachrèse en lien avec l'artéfact).

Pour ce faire, chaque itération du protocole semi-expérimental donne lieu à un même paramétrage de Moodle Atelier. En effet, ce paramétrage est globalement similaire d'une itération à l'autre, du fait du cadre méthodologique proposé par le chercheur-intervenant, avec des modifications marginales qui permettent toutefois à chaque enseignante partenaire d'adapter dans une certaine mesure le dispositif à ses propres objectifs didactiques et pédagogiques, *e.g.* en lien avec les prescriptions institutionnelles (Programmes du collège).

Ces différences mineures de paramétrage sont donc liées au caractère écologique de cette recherche-intervention, ainsi qu'à sa dimension collaborative ; elles ont également évolué à mesure que les enseignantes partenaires se sont emparées du dispositif pédagogique numérique, depuis les premiers pas de l'année N, fortement accompagnés par le chercheur-intervenant, jusqu'à la quasi-autonomie de l'année N+1.

Cependant cette personnalisation demeure périphérique : le noyau du protocole reste globalement identique tout au long de la période semi-expérimentale. Le Tableau 2 récapitule ces différentes opérations d'instrumentalisation/catachrèse, en les replaçant dans le contexte du protocole global.

PREPARATION DE L'ACTIVITE PAR L'ENSEIGNANTE ET LES ELEVES	
Consigne d'écriture et compétence-cible en étude de la langue	
Critères de réussite : méso- ou micro-compétences (approche par compétences)	

Temps de réflexion docimologique ²⁹ avec les élèves	
Descriptifs co-construits pour chaque critère de réussite (grille de coévaluation)	
CONFIGURATION DE L'ATELIER EN AMONT	
Informations générales	
Réglages de remise des travaux	
Feedback entre élèves : requis ? optionnel ?	
Travaux exemplaires ?	
Modalités d'évaluation (auto-évaluation ?)	
Attribuer localement le rôle « Etudiant » à l'enseignant	Catachrèse (1 ^{ère} opération)
Supprimer localement la permission : « Voir les auteurs » pour les « Etudiants »	Instrumentalisation (schème d'utilisation : inter-évaluation en double aveugle)
Pondération : affecter coeff. 16 pour évaluations par l'enseignant	Catachrèse (2 ^{ème} opération)
Réglages d'évaluation : par critères (note sur 20 ? sur 100 ?)	
Déterminer le nb d'évaluations par élève (3 ? 4 ? avec / sans auto-évaluation ?)	
Préparer le formulaire d'évaluation : Formulation des critères, nb de critères	
MANIPULATIONS SUR L'ATELIER EN AVAL	
Calibrer le degré d'exigence de l'algorithme : indulgent ? strict ?	
Calculer les notes	

Tableau 2 – Opérations d'instrumentalisation pour la catachrèse « coévaluation » du protocole semi-expérimental

²⁹ Docimologie : Science des examens et des concours, étude de la qualité et de la validité des différents systèmes de notation scolaire et de contrôle des connaissances (source : cnrtl.fr).

Une fois les opérations de paramétrage réalisées, les élèves pouvaient produire leur écrit et le téléverser de façon anonyme dans Moodle Atelier. On peut alors pratiquer la *coévaluation individualisée en double aveugle*, directement dans le logiciel, selon l’approche par compétences détaillée dans le formulaire d’évaluation – *i.e.* la grille récapitulant les critères de réussite, préalablement co-construite en classe par l’enseignant et les apprenants. En outre, les coévaluateurs avaient la possibilité – ou l’obligation, selon le paramétrage choisi par l’enseignante partenaire – d’y adjoindre un écrit commentatif à destination de l’élève évalué (Figure 32).

Note : 58 sur 100

Formulaire d'évaluation ▾

Grille

Critères	Niveaux			
- le texte contient des informations : date de création, acteur, décor, thème, intrigue, metteur en scène, dramaturge... > décrire ce qu'on voit > employer le vocabulaire spécifique au théâtre	<input type="radio"/> Insuffisant	<input checked="" type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
- formuler un jugement sur ce qu'on voit (adjectif, verbe d'opinion...)	<input type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input checked="" type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
- forme du texte : paragraphe, texte organisé, connecteurs...	<input type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input checked="" type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
- la critique doit montrer qu'on a bien compris le texte.	<input type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input checked="" type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant

Feedback général ▾

C'est satisfaisant sauf que au debut c'est marqué qu'on ne connais pas les acteurs alors que c'est faux on les connais.A part ça le text est assez bien organisé.

Figure 32 – Exemple de grille de coévaluation avec écrit commentatif entre élèves

Le nombre de coévaluations par élève est déterminé par l’enseignante partenaire, *i.e.* deux, trois ou quatre coévaluations par élève en règle générale. Ces éléments ont déjà été précisés *supra*, en lien avec l’artéfact didactique. Toutefois, précisons ici que l’attribution des travaux se fait de manière aléatoire dans Moodle Atelier : une fois le nombre de coévaluation par élève paramétré, le logiciel prend en charge la distribution de façon automatisée. On peut ensuite rééquilibrer au besoin, en mode manuel (Figure 33).

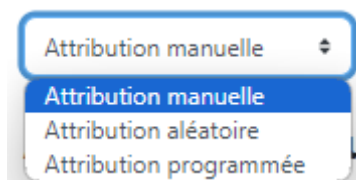


Figure 33 – Attribution des travaux dans Moodle Atelier : modes manuel et aléatoire

3.5 Durée et nombre d'itérations

La période expérimentale (année N) concerne l'année scolaire 2021-2022. Le protocole semi-expérimental a été mis en œuvre à deux reprises avec les échantillons A, A' et B, et à trois reprises avec l'échantillon C – soit un total de neuf itérations entre novembre 2021 et juin 2022.

Lors de l'année N+1 (année scolaire 2022-2023), deux expérimentations supplémentaires ont été menées en autonomie par l'enseignante B avec une nouvelle classe de 3^e (échantillon B'), en novembre et décembre. Ces expérimentations ont été réalisées en dehors du protocole semi-expérimental, toutefois elles font l'objet d'un recueil de données textuelles qui sont analysées qualitativement *infra*.

Les échantillons concernés par chaque temps de mesure ou itération hors protocole sont récapitulés chronologiquement dans la Figure 34.

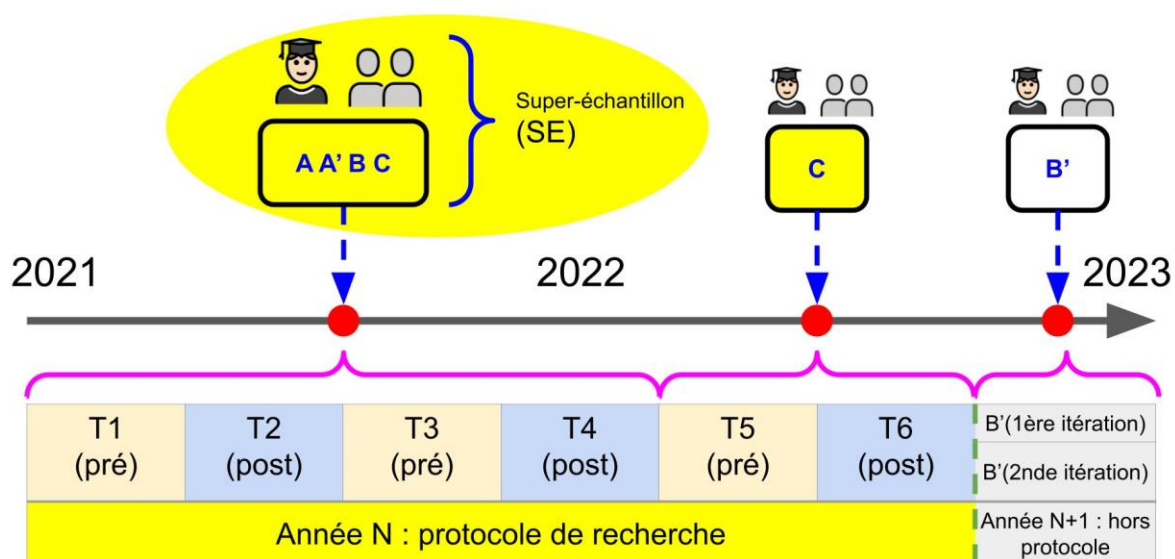


Figure 34 – Représentation chronologique de la recherche, avec temps de mesure et échantillons concernés

4 Recueil et traitement des données

Dans ce chapitre, nous présentons les outils de recueil utilisés dans la présente recherche, ainsi que les modalités de traitement des données afférentes. S'agissant d'une recherche par méthodes mixtes, les méthodes d'analyse qualitative ad hoc sont précisées in situ le cas échéant. Les méthodes d'analyse quantitative sur données numériques ou textuelles font

quant à elles l'objet d'une section dédiée, incluant notamment les approches statistiques afférentes à chacune de nos hypothèses opérationnelles.

4.1 Questionnaire en ligne

Le questionnaire en ligne dont nous nous servons pour la présente recherche revêt une importance particulière dans la mesure où :

- Il est composé de plusieurs éléments complémentaires, détaillés *infra*, e.g. :
 - l'échelle SEPE ;
 - un item portant sur le sentiment d'efficacité collective ;
 - un item portant sur la conflictualité instrumentale perçue ;
- Il est administré de manière systématique à chaque itération expérimentale selon une passation pré/post, l'objectif de cette approche pré/post systématique étant de compenser le délai potentiellement conséquent entre deux itérations du protocole ;
- Il permet de recueillir une variété de données textuelles et numériques (*i.e.* ordinales).

Nous présentons ici ce questionnaire de façon détaillée, notamment eu égard à :

- la manière dont il est structuré ;
- la nature des items qui composent ses différentes sections ;
- certains aspects statistiques afférents.

4.1.1 Codage des temps de mesure

En année N, nos temps de mesure avec tous les échantillons (A, A', B et C) correspondent à la passation pré/post du questionnaire psychométrique SEPE, auquel sont rattachés d'autres items, tous décrits *infra*.

- T1 : juste avant la 1^{ère} exposition au dispositif ;
- T2 : juste après la 1^{ère} exposition au dispositif ;
- T3 : juste avant la 2^e exposition au dispositif ;
- T4 : juste après la 2^e exposition au dispositif.

Avec l'échantillon C uniquement, deux temps de mesure supplémentaires :

- T5 : juste avant la 3^e exposition au dispositif ;
- T6 : juste après la 3^e exposition au dispositif.

En année N+1, avec l'échantillon B', il n'y a pas de passation du questionnaire, donc pas de temps de mesure codé. Les données numériques et textuelles recueillies font néanmoins l'objet d'un traitement spécifique *infra*.

4.1.2 Structure générale du questionnaire

- Sur la quarantaine d'items composant le questionnaire pré/post, 37 sont dévolus à la mesure du sentiment d'efficacité personnelle des enfants (SEPE).
- Lors de la première et de la troisième passation, conformément aux recommandations détaillées *supra*, la partie SEPE stricto sensu est précédée d'une série de 8 questions d'entraînement, *i.e.* questions de calibrage, codées « CAL », visant à familiariser les participants avec la formulation des questions des échelles SEP et SEC, ainsi qu'avec l'échelle de Likert à cinq points afférente.
- Lors de la toute première passation (*i.e.* à T1), deux items visent à mesurer les habitudes évaluatives des élèves en matière d'auto-évaluation d'une part, et d'inter-évaluation d'autre part.
- A cela s'ajoute un item unique visant à mesurer le sentiment d'efficacité collective (SEC) des élèves participants.
- Deux items visent à évaluer la conflictualité instrumentale vécue par les participants³⁰ :
 - un item générant une donnée quantitative (*i.e.* ordinale)
 - un item générant une donnée qualitative (*i.e.* textuelle)
- A T2, une question ouverte invite chaque participant à faire librement part de son retour sur expérience concernant le dispositif de coévaluation numérique.
- A T4, à la fin du questionnaire, trois questions invitent les élèves à en évaluer l'acceptabilité, l'objectif étant de renseigner le chercheur quant à la validité de ce processus de collecte de données autorapportées.

Ces différents éléments sont résumés dans le Tableau 3. Les libellés de tous les items sont par ailleurs consultables en Annexes 3. Les sections qui suivent apportent des

³⁰ La notion de *conflict instrumental* n'est bien sûr pas conscientisée par les élèves participants. La formulation des items a cependant pour but de mettre au jour cette conflictualité, en les interrogeant sur leur ressenti avec des termes simples.

compléments d'information concernant certaines parties du questionnaire. En particulier, les méthodes d'analyse qualitative sont précisées le cas échéant. Le traitement quantitatif des données numériques et textuelles est quant à lui détaillé dans une section dédiée.

PARTIES DU QUESTIONNAIRE	NOMBRE D'ITEMS	TYPE(S) DE(S) QUESTION(S)	TEMPS DE MESURE CONCERNES					
			T1	T2	T3	T4	T5	T6
Calibrage	8	Likert 5 points	8		8		8	
SEPE	37	Likert 5 points	37	37	37	37	37	37
SEC	1	Likert 5 points			1	1	1	1
Expérience préalable (auto- et inter-éval.)	2	Likert 5 points	2					
CI-EIAH	2	Likert 5 points Question ouverte		2		2		2
Retour sur expérience	1	Question ouverte		1				
Acceptabilité questionnaire	3	Likert 5 points				3		
TOTAL			47	40	46	43	46	40

Tableau 3 – Structure détaillée du questionnaire en ligne en fonction des temps de mesure

4.1.3 Construction du questionnaire SEPE

Selon une approche pré/post, nous utilisons l'échelle psychométrique SEPE pour mesurer le sentiment d'efficacité personnelle des élèves participants avant et après chacune des deux expositions au dispositif expérimental – soit pour un total de 4 temps de mesure : T1, T2, T3 et T4. A ceci s'ajoutent les temps de mesure T5 et T6 pour l'échantillon C.

4.1.3.1 Sous-échelles du SEPE

Parmi les sept dimensions de l'échelle SEPE, trois retiennent spécifiquement notre attention en raison de leur pertinence pour tester trois de nos hypothèses opérationnelles (HO 1, HO 2a et HO 3) :

- L'auto-efficacité perçue pour la réussite scolaire ou SEP(scol), qui mesure les représentations des élèves en termes de *capacité à maîtriser des compétences dans différentes disciplines scolaires*. Cette sous-échelle se rapporte à la réussite scolaire en général ; toutefois sur les sept items qui la composent deux portent spécifiquement sur l'apprentissage du français. Nous l'utilisons donc pour tester

HO 1, *i.e.* « La coévaluation numérique favorise l'amélioration des compétences en français ».

- L'auto-efficacité perçue pour l'autorégulation des apprentissages ou SEP(srl), qui évalue les représentations des élèves en termes de capacité à *structurer leur environnement pour le rendre propice aux apprentissages, ainsi qu'à planifier et organiser des activités académiques* ; cet indicateur permet de tester HO 2a, *i.e.* « La coévaluation numérique favorise l'autorégulation des apprentissages ».
- L'auto-efficacité perçue pour les apprentissages coopératifs ou SEP(coop), qui mesure les représentations des élèves en termes de *capacité à initier et cultiver des relations sociales et à gérer les conflits interpersonnels* ; cet indicateur permet de tester HO 3, *i.e.* « La coévaluation numérique favorise les apprentissages coopératifs ».

Les libellés des différents items qui composent ces trois sous-échelles, ainsi que leur nombre, sont précisés en Annexes 3, respectivement dans le Tableau 65, le Tableau 66 et le Tableau 67.

4.1.3.2 Respect des recommandations de l'auteur

Notre échelle SEPE comporte 37 items, conformément aux recommandations de l'auteur (Bandura, 2006), à l'instar de l'étude conduite par (Pastorelli *et al.*, 2001) auprès d'écoliers d'un âge moyen de 12,5 ans – âge équivalent à celui des participants de notre propre étude.

Nous sommes par ailleurs particulièrement attentif aux recommandations de l'auteur concernant :

- la pertinence et la cohérence des dimensions retenues ;
- le nombre d'items suffisamment important pour étayer statistiquement les résultats ;
- le calibrage préliminaire du test grâce à des questions d'entraînement posées aux participants ;
- l'utilisation de pictogrammes – en l'occurrence ici des étoiles – plutôt que de chiffres, en raison de la jeunesse du public.

4.1.3.3 Processus utilisé pour la traduction de l'échelle en français

Les 37 items originellement formulés en langue anglaise font l'objet d'un processus de traduction et rétro-traduction en double aveugle par deux locuteurs bilingues indépendants,

avant une harmonisation permettant d'aboutir à une version de travail consensuelle. Cette version est ensuite prétestée pour en vérifier la validité interne au moyen de l'alpha de Cronbach et éliminer ou reformuler les items ambigus.

4.1.3.4 Echelle de Likert à cinq points

Comme dans l'étude de (Pastorelli *et al.*, 2001), les réponses aux questions sont récoltées au moyen d'une échelle de Likert à cinq points. Pour chacun des items, les 5 points de l'échelle de Likert sont représentés sous la forme d'étoiles, accompagnées d'un descriptif pour aider les participants à en interpréter la signification (Figure 35). Ce codage génère les données ordinales suivantes :

1. Je ne m'en sens pas du tout capable.
2. Je m'en sens peu capable.
3. Je m'en sens moyennement capable.
4. Je m'en sens capable.
5. Je m'en sens tout à fait capable.

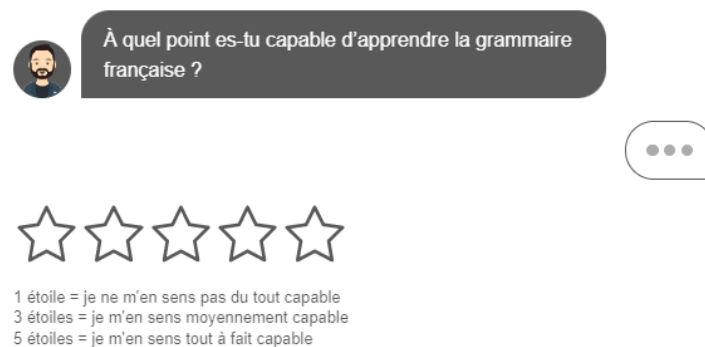


Figure 35 – Exemple d'item avec échelle de Likert et descriptif (interface participants)

Toutes les questions fermées de notre questionnaire en ligne appellent une réponse via une échelle de Likert à cinq points dont le fonctionnement est analogue à celle figurant ci-dessus. Bien que textuelles à l'origine, ces données satisfont à des critères précis, justifiant leur transposition sous forme de données ordinales et leur traitement statistique sous forme de données à intervalle. Il s'agit en somme d'appliquer une approche quantitative à des données initialement textuelles, mais converties en données numériques. C'est la raison pour laquelle les détails et la justification de ce traitement statistique particulier sont fournis *infra*, dans une section dédiée au traitement des données numériques.

4.1.3.5 Items de calibrage

Aux temps de mesure T1, T3 et T5, conformément aux recommandations de (Bandura, 2006), nous faisons précéder les 37 items de l'échelle SEPE d'une série de huit items dits de « calibrage », *i.e.* des items destinés à familiariser les participants à notre échelle de Likert, à sa signification et à sa logique de fonctionnement (Figure 36). L'objectif est également d'optimiser, et dans une certaine mesure de vérifier, la validité de ces données autorapportées. Concrètement, ces huit items correspondent à des tâches de difficulté croissante, *i.e.* de CAL 1 (tâche la plus facile) à CAL 8 (tâche la plus difficile). Selon la logique de ces items de Likert, le sentiment d'efficacité personnel est censé diminuer à mesure que la difficulté augmente.

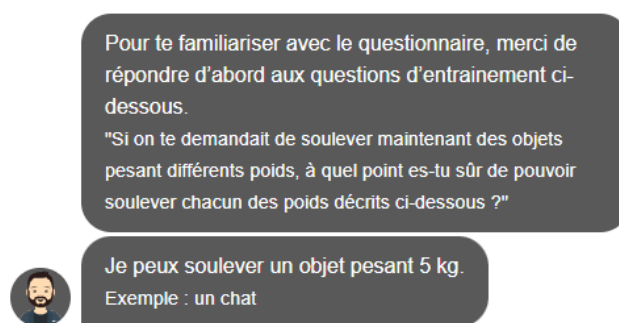


Figure 36 – Exemple d'item de calibrage

4.1.3.6 Caractère non indispensable des analyses factorielles ici

Dans la mesure où nous utilisons une échelle déjà validée par des analyses factorielles dans des études à grande échelle, à l'instar de celles mentionnées *supra* (Bandura, 2006; Bandura *et al.*, 1996; Foster *et al.*, 2014; Pastorelli *et al.*, 2001), il n'est pas nécessaire de procéder ici à une analyse factorielle exploratoire pour identifier de potentiels facteurs latents. Par conséquent, une analyse en composantes principales et une analyse factorielle confirmatoire ne sont pas non plus pertinentes (Navarro & Foxcroft, 2022).

Néanmoins, afin de valider le comportement de l'échelle avec notre propre échantillon, nous procédons à une analyse de fiabilité appréciant la cohérence interne à l'aide des coefficients alpha de (Cronbach, 1951) et oméga de (McDonald, 2013, 2014). Les résultats de ces analyses de cohérence interne sont présentés *infra*, dans notre partie « Résultats » (partie IV).

4.1.4 Questions portant sur les habitudes évaluatives

Au début du questionnaire en ligne, uniquement à T1, pour renseigner le profil de chaque élève participant relativement à l'existence d'une expérience préalable avec les

artéfacts pédagogiques *auto-évaluation* et *évaluation par les pairs*, nous intégrons deux items ad hoc. Les données sont collectées via une échelle de Likert à cinq points. Elles génèrent deux valeurs brutes indépendantes, auxquelles on applique le traitement statistique détaillé *infra*.

Précisons qu'aucun item ne porte sur la coévaluation car cette pratique encore émergente peut constituer un concept flou pour des collégiens de cycle 4 ; tandis que parler d'évaluation entre pairs semble pertinent dans la mesure où, au départ, cette activité constitue pour les élèves la partie visible de la coévaluation – l'autre composante relevant plutôt de l'enseignant, tout du moins dans un premier temps.

Par ailleurs, la présence d'un item portant sur l'expérience préalable en matière d'auto-évaluation se trouve justifiée par le fait que Moodle Atelier permet d'intégrer cette pratique au dispositif de coévaluation. En outre, nous avons vu dans notre état de l'art que l'auto-évaluation présentait plusieurs caractéristiques communes avec l'inter-évaluation et la coévaluation.

4.1.5 Échelle SEC à item unique (indice agrégatif personnel)

Pour tester HO 3 (*i.e.* « La coévaluation numérique favorise les apprentissages coopératifs ») en effectuant des recoupements entre variables opérationnelles liées à la dimension collective et/ou coopérative des apprentissages, nous avons également opté pour une échelle mesurant le sentiment d'efficacité collective des élèves (1 item), basée sur l'indice agrégatif personnel selon (Bandura, 2000).

En effet, le dispositif de coévaluation numérique que nous expérimentons ici comporte une composante coopérative significative, bien qu'il s'agisse d'une modalité hybride, *i.e.* :

- en présentiel en salle informatique ;
- mais par ordinateurs interposés ;
- et avec l'asynchronie qu'impliquent les différences inter-individuelles en termes de rythme de travail.

Mesurer le sentiment d'efficacité collective dans ce contexte semble pertinent, on peut donc envisager des emprunts à (Cosnefroy & Jézégou, 2013) ou (Wang & Lin, 2007).

Il ne s'avère toutefois pas pertinent de transposer les items de l'échelle de Cosnefroy & Jézégou (2013) à notre étude dans la mesure où chez ces auteurs il s'agissait d'un travail *de* groupe, *i.e.* une activité à forte interdépendance. Tandis que l'activité de coévaluation numérique individualisée que nous expérimentons ici est un travail *en* groupe, au sens où l'activité individuelle des apprenants est coordonnée au service d'un but commun (*i.e.* la

coévaluation, pour aider chacun à progresser), mais avec une faible interdépendance entre les acteurs.

De même, dans le cas de l'échelle utilisée par (Wang & Lin, 2007, p. 2260), le niveau groupal prédomine, avec une forte interdépendance entre les membres : on cherche en effet à mesurer la « capacité du groupe, considéré comme un tout, à performer ».

Or dans le cas du dispositif de coévaluation formative individualisée qui fait l'objet de la présente recherche, il s'avère que l'interdépendance entre les apprenants est basse. Chacun effectue son opération d'inter-évaluation de son côté, tout en ayant conscience de contribuer à un processus collectif dont il ou elle bénéficie au même titre que les autres : « J'évalue la production écrite de plusieurs pairs, et plusieurs pairs évaluent la mienne ».

Soulignons que, bien qu'elle soit basse, l'interdépendance entre les membres du groupe n'est pas inexistante. Le dispositif implique en effet – lorsque l'enseignant a choisi de rendre cela obligatoire – la production d'un écrit commentatif, *i.e.* une appréciation rédigée à destination du pair dont on vient d'évaluer la production écrite.

Autre élément d'interdépendance : tout élève participant est en quelque sorte « comptable » de son inter-évaluation ; en effet, par l'entremise de l'artéfact numérique, cette dernière sera comparée, compétence par compétence, avec l'hétéro-évaluation de référence, effectuée par l'enseignant, ce qui génèrera un « score de coévaluateur ».

On le voit : le concept-même de coévaluation implique plusieurs formes d'interdépendance. Toutefois, ce niveau d'interdépendance reste relativement faible : on est ici plus proche du fonctionnement collectif d'une équipe de gymnastique que de celui d'une équipe de football – pour reprendre l'analogie de (Bandura, 2000) évoquée *supra*. C'est pourquoi, dans le cas qui nous intéresse, une mesure pertinente du sentiment d'efficacité collective (SEC) relèvera de l'indice agrégatif personnel (IAP).

Les échelles des auteurs mentionnés précédemment (Cosnefroy & Jézégou, 2013; Wang & Lin, 2007) correspondent à une approche holistique et ne peuvent pas être utilisées pour la présente recherche. Cela implique de construire notre propre échelle, selon l'approche de l'IAP, pour mesurer le *sentiment d'efficacité de chaque membre du groupe* concernant sa *capacité personnelle à accomplir les fonctions spécifiques qui lui incombent* au sein de celui-ci ; avant d'agrèger ces mesures pour obtenir une mesure du sentiment d'efficacité collective selon l'indice agrégatif personnel (SEC-IAP).

Au temps de mesure T2 et T4 du protocole, nous intégrons donc un item unique visant à mesurer le SEC-IAP de chaque participant. S'agissant d'une échelle de Likert

unidimensionnelle à item unique, il n'est pas nécessaire de pratiquer une AFE³¹, ni une AFC³², ni une ACP³³, ni même une analyse de fiabilité (cohérence interne).

L'item en question est formulé comme suit (Figure 37) :

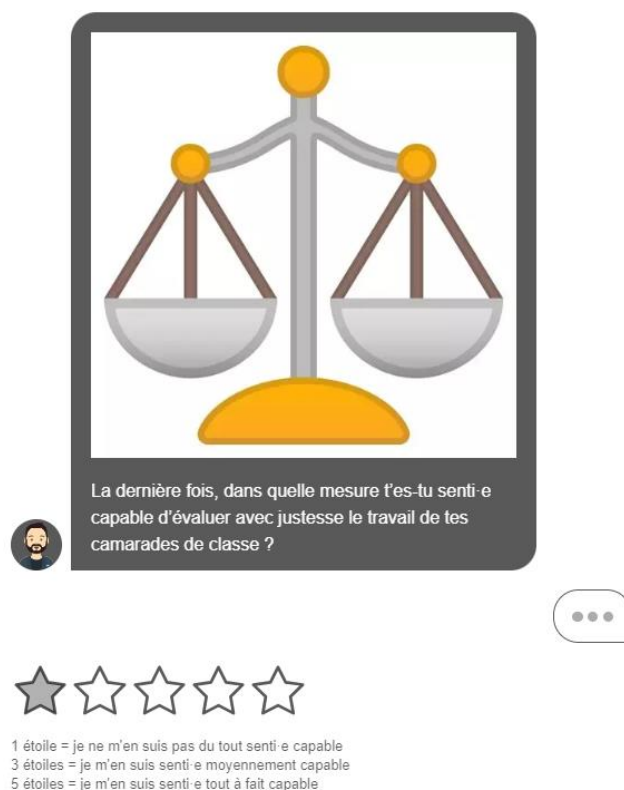


Figure 37 – Item unique mesurant le SEC-IAP dans le questionnaire en ligne

A l'instar de l'ensemble des questions fermées du questionnaire en ligne, cet item génère une donnée textuelle, en l'occurrence la réponse de l'élève participant sur l'échelle de Likert à cinq points figurant ci-dessus. Pour les besoins de l'analyse statistique, cette donnée textuelle est transposée sous forme de donnée ordinale, puis cette valeur brute est traitée quantitativement en tant que donnée à intervalle, selon les modalités détaillées *infra* dans la section « Traitement des données numériques ».

4.1.6 Questions portant sur la genèse instrumentale

A T2 et à T4, à la fin du questionnaire en ligne, trois items³⁴ visent à nous renseigner concernant la genèse instrumentale du dispositif, du point de vue des élèves, c'est-à-dire à tester

³¹ AFE : Analyse factorielle exploratoire.

³² AFC : Analyse factorielle confirmatoire.

³³ ACP : Analyse en composantes principales.

³⁴ Le troisième item n'est proposé qu'à T2 (voir *supra*).

l'hypothèse HO 4 – c'est-à-dire « La coévaluation numérique génère des conflits instrumentaux et permet de les réguler ». Sans faire explicitement référence à l'approche instrumentale stricto sensu, il s'agit d'utiliser des mots simples pour recueillir le ressenti spontané des élèves relativement aux conflits instrumentaux rencontrés et aux régulations instrumentales éventuellement mises en œuvre.

4.1.6.1 Item 1

La première question – c'est-à-dire « La coévaluation avec Moodle Atelier t'a-t-elle paru difficile ? » (item 1) – a pour but de nous renseigner sur le degré de conflictualité instrumentale ressenti par les élèves, lors des situations d'enseignement/apprentissage mises en œuvre ici. L'item 1 consiste en une question fermée, cadrée par l'échelle de Likert ci-dessous (Figure 38) :



C'est presque fini !
Avant de se quitter, j'aimerais juste avoir ton avis sur l'expérience que tu viens de vivre.

La coévaluation avec l'atelier Moodle t'a-t-elle parue plutôt facile ou difficile ?

★ ★ ★ ★ ★

1 étoile = très difficile
3 étoiles = facile et difficile à la fois
5 étoiles = très facile

Figure 38 – Question fermée (item 1) portant sur la genèse instrumentale du dispositif expérimental

Les données textuelles obtenues sont transposées en données ordinales puis analysées quantitativement en tant que données à intervalle. Puisqu'il s'agit d'un item unique, nous utilisons les valeurs brutes. Sur échantillons appariés, nous comparons ces scores au moyen :

- du test de Wilcoxon :
 - sur l'intervalle T2-T4 ;
 - ainsi que sur l'intervalle T4-T6 (échantillon C uniquement) ;
- du test de Friedman sur l'intervalle T2-T6 (échantillon C uniquement).

4.1.6.2 Item 2

La deuxième question vise à compléter ces informations quantitatives, en générant des données textuelles sous la forme d'un retour sur expérience semidirigé répondant à la question ouverte « Pourquoi ? » (item 2), et invitant à fournir un (ou plusieurs) exemple(s) concret(s). Les données textuelles ainsi recueillies sont analysées au moyen d'une analyse catégorielle de contenus (c'est-à-dire une analyse thématique). Cette analyse se fait ici en deux temps : une analyse préliminaire permet d'identifier les principales thématiques abordées, puis une seconde analyse catégorielle permet de les analyser artéfact par artéfact, eu égard au cadre théorique de référence ici, c'est-à-dire l'approche instrumentale en environnement informatisé pour l'apprentissage humain (EIAH). Pour en faciliter l'intelligibilité, les verbatim d'élèves illustrant les différentes thématiques ont été redressés/toiletés (*e.g.* ponctuation, graphies, organisation paragraphique des réponses). La Figure 39 est une capture d'écran illustrant cet item du questionnaire en ligne.

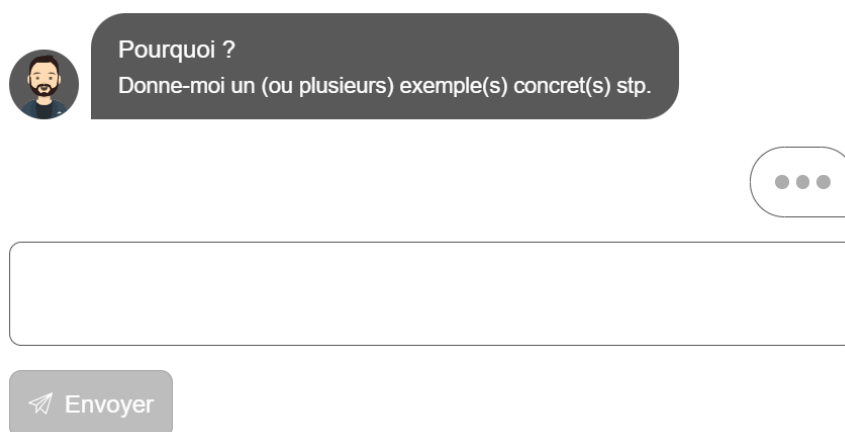


Figure 39 – Question ouverte (item 2) portant sur la genèse instrumentale du dispositif expérimental

4.1.6.3 Item 3

Dans le questionnaire en ligne, un troisième item destiné à tester l'hypothèse HO 4 – c'est-à-dire « La coévaluation numérique génère des conflits instrumentaux et permet de les réguler » – consiste en une question libre (voir capture d'écran, Figure 40). En raison de son caractère potentiellement redondant vis-à-vis de l'item 2, l'item 3 n'est proposé qu'une seule fois, au temps de mesure T2. Les données textuelles ainsi recueillies sont analysées au moyen d'une analyse catégorielle de contenus. Là encore, l'analyse se fait en deux temps : une analyse préliminaire permet d'identifier les principales thématiques abordées, puis une seconde analyse catégorielle permet de les analyser artéfact par artéfact, eu égard au cadre théorique de référence ici, c'est-à-dire l'approche instrumentale en EIAH. Toujours pour en faciliter

l'intelligibilité, les verbatim d'élèves illustrant les différentes thématiques ont été redressés/toiletés (e.g. ponctuation, graphies, organisation paragraphique des réponses).

Pour finir, as-tu des remarques ou des commentaires à faire ? Ici même les critiques sont les bienvenues ;)

Par exemple, tes commentaires peuvent porter sur cette expérience de coévaluation avec Moodle, ou même sur ce questionnaire.



Figure 40 – Question libre (item 3) portant sur la genèse instrumentale du dispositif expérimental

4.1.7 Questions relatives à l'acceptabilité du questionnaire

Compte tenu de la relative longueur du questionnaire et du public pré-adolescent ciblé, trois questions ad hoc invitent les élèves à en évaluer l'ergonomie, à T4 – c'est-à-dire lors du dernier temps de mesure –, à la toute fin du questionnaire. L'objectif est de vérifier l'acceptabilité du questionnaire du point de vue des participants, élément susceptible de nous renseigner quant à la fiabilité de nos données autorapportées. Ces trois questions (Tableau 4) visent à évaluer respectivement :

- La facilité ou difficulté de compréhension du questionnaire, eu égard à la formulation des questions (item 1) ;
- L'attrait exercé par ce questionnaire en ligne, compte tenu des éléments de ludification liés à Tripetto® ;
- La facilité ou difficulté de prise en main du questionnaire par les élèves participants (item 3).

1	Pour toi, ce questionnaire était-il plutôt facile ou difficile à comprendre ?
2	Pour toi, l'apparence de ce questionnaire était-elle plutôt agréable ou désagréable ?
3	Pour toi, ce questionnaire était-il plutôt pratique ou peu pratique ?

Tableau 4 – Les trois items visant à mesurer l'ergonomie du questionnaire selon les participants

Pour chaque item, une échelle de Likert à cinq points symbolisés par des étoiles permettait aux élèves participants d'exprimer leur avis. Un descriptif visait à guider les participants concernant l'interprétation de cette échelle de Likert (voir Figure 41).



Figure 41 – Exemple avec échelle de Likert à cinq points et descriptif

Pour chaque item, les échelles de Likert étaient décrites comme suit :

- Pour l’item 1 (compréhension/formulation des questions) :
 - une étoile = très difficile à comprendre
 - trois étoiles = un peu difficile / pas facile
 - cinq étoiles = très facile à comprendre
- Pour l’item 2 (esthétique/aspect ludique) :
 - une étoile = très désagréable
 - trois étoiles = un peu désagréable / peu agréable
 - cinq étoiles = très agréable
- Pour l’item 3 (ergonomie/prise en main) :
 - une étoile = pas du tout pratique
 - trois étoiles = peu pratique
 - cinq étoiles = très pratique

Les modalités de traitement des données ainsi obtenues font l’objet d’une description d’ensemble, *infra*.

4.2 Traitement des données numériques

Pour analyser les données textuelles recueillies via le questionnaire en ligne, nous utilisons une approche quantitative. Nous présentons ici plusieurs principes statistiques qui ont structuré l’analyse.

4.2.1 Traitement statistique des variables ordinales en tant que données à intervalle

Pour commencer, nous avons tenu compte du problème d'interprétation lié au fait de traiter des données ordinales comme des données continues – en l'occurrence, des données à intervalle –, dans la mesure où cela reviendrait par exemple à considérer à tort l'écart ressenti entre « Je ne m'en sens pas du tout capable » et « Je m'en sens peu capable » comme équivalent à celui ressenti entre les degrés suivants de l'échelle de Likert (Cohen *et al.*, 2002).

Toutefois, le fait de traiter des données ordinales issues d'une échelle de Likert comme des données continues à échelle d'intervalle³⁵ – autrement dit *interval scale data* (Ary *et al.*, 2010) – peut être considéré comme légitime, à condition que les échelles utilisées soient des agrégats d'au moins 4 items mesurant effectivement le même facteur (Boone & Boone, 2012).

Or l'échelle SEPE utilisée ici est une échelle déjà bien établie, pour laquelle des analyses factorielles, comme celles mentionnées dans l'état de l'art, ont déjà été menées à grande échelle. En outre, il est possible d'établir au moyen d'une analyse de cohérence interne – *e.g.* alpha de Cronbach, oméga de McDonald – que l'échelle SEPE et les sous-échelles afférentes SEP(scol), SEP(sr1) et SEP(coop) constituent effectivement des échelles de Likert fiables, avec des coefficients de corrélation satisfaisants et aucun effet de redondance constaté. Dans ce but, les résultats de nos analyses de cohérence interne sont présentés *infra*.

Pour les besoins de l'analyse statistique, il est donc possible de traiter des données ordinales, issues d'échelles de Likert *avérées*, au sens de (Boone & Boone, 2012), comme des données continues à échelle d'intervalle – autrement dit *interval scale data* (Ary *et al.*, 2010). Cependant il ne s'agit pas d'interpréter *in fine* les quantités ou proportions ainsi constatées de façon littérale. Relativement aux variables qualitatives mesurées par les items de chaque échelle ou sous-échelle, il s'agit plutôt de considérer simplement les implications des différences significatives observées à la hausse ou à la baisse, en s'inspirant de l'approche clinique, comme le suggèrent (Portney & Watkins, 2009; Sheldon *et al.*, 1996). C'est selon cette logique que, dans la présente recherche, nous avons transposé les données textuelles

³⁵ Les variables à échelle d'intervalle sont « similaires aux variables ordinales, sauf que les intervalles entre les valeurs de la variable à échelle d'intervalle sont espacés de manière égale. Un bon exemple d'une échelle d'intervalle est l'échelle des degrés Fahrenheit utilisée pour mesurer la température. Avec l'échelle Fahrenheit, la différence entre 70° et 75° est égale à la différence entre 80° et 85°: les unités de mesure sont égales pour toute l'étendue de l'échelle » (Ali & Bhaskar, 2016, p. 663, traduction libre).

issues des échelles de Likert en données numériques, c'est-à-dire en données ordinales puis en données à intervalle.

Il s'avère enfin que les tests non-paramétriques constituent un choix à privilégier pour analyser quantitativement ce type de données, dans la mesure où ils tendent à générer des statistiques inférentielles basées sur les *rangs* – et non pas basés sur des construits arithmétiques nécessitant des données avec échelle de rapport (« *ratio scale data* ») ou échelle d'intervalle (« *interval scale data* »), comme c'est le cas pour les tests paramétriques (Jamieson, 2004).

4.2.2 Difficulté d'établir une distribution normale avec de petits échantillons ($N \leq 30$)

Un autre aspect important de notre étude peut inciter à opter pour des tests non-paramétriques. En effet, si l'ensemble des participants (tous échantillons confondus, de T1 à T4), s'élève à ($N = 80$), les échantillons *appariés* sur lesquels nous devons nous baser pour nos analyses inférentielles sont souvent de taille largement moindre ($4 \leq N \leq 62$).

Or pour vérifier si des données sont normalement distribuées – conditions sine qua non pour pratiquer des tests paramétriques –, le test de Shapiro-Wilk est usuellement recommandé avec les échantillons de petite taille ($N < 50$) – et même, selon (Shapiro & Wilk, 1965, p. 610) eux-mêmes, avec les échantillons de très petite taille ($N < 20$).

Toutefois, (Le Boedec, 2016) a montré dans une étude randomisée à grande échelle que le test de Shapiro-Wilk se révèle peu significatif avec des échantillons de très petite taille ($N = 30$). Plus précisément, (Souza *et al.*, 2023, p. 8) ont récemment montré qu'une population de ($17 \leq N \leq 20$) constitue un seuil en deçà duquel les résultats peuvent présenter une variabilité significativement accrue et, de ce fait, être la source d'erreurs d'analyse de la normalité via le test de Shapiro-Wilk. Ces possibles biais méthodologiques peuvent être contournés en pratiquant exclusivement des tests non-paramétriques.

4.2.3 Priorité donnée aux comparaisons intragroupes

Compte tenu de la situation naturelle, impliquant la présence de nombreux paramètres différents d'un échantillon à l'autre – *e.g.* enseignants différents, niveaux différents (5^e, 4^e ou 3^e), types d'établissements différents (rural, périurbain, urbain en réseau d'éducation prioritaire) – nous privilégions ici une analyse séparée des données de chaque échantillon,

c'est-à-dire une approche comparative intragroupes entre n temps de mesure, *e.g.* SEC-IAP de l'échantillon B à T3 vs. SEC-IAP du même échantillon (apparié) à T4.

4.2.4 Approche complémentaire : « super-échantillon » (SE)

En complément des comparaisons intragroupes mentionnées *supra*, des analyses complémentaires sont menées sur l'ensemble des échantillons, considérés comme un seul échantillon global et dénommé ci-après « super-échantillon » (SE). Ce choix est légitimé par la relative homogénéité, à la fois du protocole semi-expérimental co-construit avec chaque enseignante partenaire, et de la population participante – c'est-à-dire des collégiens de cycle 4, tous scolarisés en établissement public local d'enseignement en France métropolitaine.

4.2.5 Choix des tests non-paramétriques pour cette étude

4.2.5.1 Une alternative adaptée

A l'instar de (Le Boedec, 2016; Souza *et al.*, 2023), (Grech & Calleja, 2018) affirment que la normalité ou non-normalité de la distribution des données est difficile voire impossible à établir avec certitude pour des échantillons de très petite taille, quels que soient la méthode d'interprétation graphique (*e.g.* diagramme Q-Q) ou le type de test utilisés (*e.g.* Shapiro-Wilk). Compte tenu de l'impossibilité de générer des données inférentielles fiables au moyen de tests paramétriques dans ce contexte, les tests non-paramétriques peuvent être considérés comme une alternative adaptée (Harpe, 2015; Jamieson, 2004; Likert, 1932; Norman, 2010), surtout avec des échantillons de petite taille (Pett, 2015).

4.2.5.2 Avantages et inconvénients de l'approche non-paramétrique

Les tests non-paramétriques présentent l'inconvénient d'être moins puissants statistiquement, c'est-à-dire qu'ils peuvent se révéler incapables de rejeter une hypothèse nulle qui s'avère fautive (St. Laurent & Turk, 2013; Zimmerman & Zumbo, 1993), et être ainsi la source de « faux-négatifs » (erreurs de type II). Néanmoins, ce type de tests semble adapté pour générer des données inférentielles à partir de variables ordinales issues d'échelles de Likert dans le cadre de designs longitudinaux, à l'instar de la présente recherche (Sheldon *et al.*, 1996).

Pour analyser nos données à intervalle issues des échelles de Likert, ainsi que nos données numériques recueillies sous forme de valeurs brutes (*e.g.* les scores de coévaluateurs), nous pratiquons donc des tests non-paramétriques.

4.2.5.3 Appariement des échantillons

Compte tenu de la dimension longitudinale de l'étude, ces analyses inférentielles sont réalisées sur échantillons appariés. En effet, pour qu'une comparaison intragroupe entre deux temps de mesure soit valide – *e.g.* pour l'échantillon B entre T3 et T4 –, l'échantillon concerné doit avoir la même composition, c'est-à-dire les mêmes élèves, d'un temps de mesure à l'autre.

4.2.5.4 Tests choisis

4.2.5.4.1 Corrélation de Spearman

Concrètement, nous utilisons la corrélation de Spearman lorsque nous cherchons à déterminer si deux variables sont corrélées. Il s'agit en effet d'une alternative non-paramétrique à la corrélation de Pearson, qui présente l'avantage d'être particulièrement robuste statistiquement, c'est-à-dire le risque que d'éventuelles données aberrantes aient un impact significatif sur sa valeur est limité.

Relativement au coefficient de corrélation *rho* de Spearman, une corrélation positive parfaite ($r = 1$) signifie que l'une des deux variables est une fonction monotone parfaite de l'autre. En d'autres termes, on se trouve en présence d'une variation simultanée des deux variables dans le même sens. A la différence du test de Pearson, on ne peut pas établir avec le test de Spearman qu'une corrélation est parfaitement linéaire ; toutefois une corrélation positive parfaite, établie avec ce dernier, signifie une correspondance totale entre les rangs.

4.2.5.4.2 Test des rangs signés de Wilcoxon

Pour comparer deux échantillons appariés, nous utilisons le test des rangs signés de Wilcoxon. Si une différence significative est détectée, le coefficient de corrélation entre rangs bisériés ($-1 < r < 1$) peut servir d'estimateur de taille d'effet, c'est-à-dire nous renseigner sur l'ampleur de la hausse ou de la baisse constatée.

4.2.5.4.3 Test de Friedman

Nous utilisons enfin le test de Friedman (ANOVA pour mesures répétées) lorsque nous avons trois échantillons appariés (ou plus) à comparer, afin de déterminer une différence significative entre ces échantillons. En cas de différence significative, le test de Durbin-Conover peut être mobilisé en analyse post-hoc – pour affiner notre comparaison entre échantillons, pair-à-pair.

4.2.6 Utilisation de l'écart interquartile

Ponctuellement, nous utilisons l'écart interquartile (EI). Dans un jeu de données, de part et d'autre de la médiane (ou deuxième quartile), le troisième quartile (Q_3) est la donnée en-deçà de laquelle se situent les 75 % inférieurs du jeu de données ; tandis que le premier quartile (Q_1) est la donnée en-deçà de laquelle se situent les 25 % inférieurs. L'écart interquartile ($EI = Q_3 - Q_1$) renseigne sur la dispersion du jeu de données : plus l'écart interquartile est élevé, plus les données sont dispersées. L'EI est considéré comme un indicateur statistique robuste (Vinutha *et al.*, 2018; Yin, 2015).

4.2.7 Utilisation des valeurs médianes et moyennes

4.2.7.1 Nécessité d'utiliser un indicateur de tendance centrale

Compte tenu de la dimension longitudinale de la présente recherche, il s'agit d'envisager la manière dont un indicateur de tendance centrale – *e.g.* médiane, moyenne – peut être utilisé pour rendre compte, sous la forme d'un seul nombre le plus représentatif possible, des 37 données à intervalle que chaque élève participant génère à chaque passation de l'échelle SEPE. Dans une moindre mesure, la même question se pose pour les sous-échelles SEP(scol), SEP(srl) et SEP(coop).

4.2.7.2 Arguments en faveur de l'utilisation de la médiane

De prime abord, pour des tests non-paramétriques pratiqués sur nos échantillons de (très) petite taille, l'utilisation des valeurs médianes semblait la mieux adaptée (Grech & Calleja, 2018; Harris *et al.*, 2008; Sheskin, 2003). En effet, la médiane est généralement considérée comme un meilleur indicateur de tendance centrale que la moyenne dans ce contexte (Gardner & Martin, 2007; Sullivan & Artino Jr, 2013), notamment parce que les manipulations arithmétiques requises pour calculer la moyenne (et l'écart-type) sont considérées comme inappropriées pour des données ordinales (Blaikie, 2003, repris par Jamieson, 2004). En outre, selon (Navarro & Foxcroft, 2022, p. 64), il y a adéquation entre le mode de calcul de la médiane et la situation qui s'applique dans le cas de variables ordinales, dans la mesure où la médiane met l'accent sur l'ordonnement dans la distribution des valeurs (rangs), mais n'est pas déterminée par les valeurs quantitatives stricto sensu.

4.2.7.3 Argument en défaveur de l'utilisation de la médiane

L'argumentation qui précède peut toutefois être réfutée. En effet, (Sheldon *et al.*, 1996) recommandent une certaine prudence interprétative vis-à-vis des résultats ainsi obtenus, compte tenu du caractère potentiellement arbitraire des valeurs assignées aux rangs via le calcul de la médiane. Selon eux, étant donné que la médiane est la valeur centrale au sein d'une distribution ordonnée sous forme de rang – et indique de ce fait une position, non pas une quantité –, elle n'est pas impactée par la valeur quantitative des scores (Portney & Watkins, 2009).

4.2.7.4 Argument en faveur de l'utilisation de la moyenne

Nous avons vu *supra* que nos données ordinales satisfont aux critères énoncés par (Boone & Boone, 2012) pour être traitées statistiquement comme des données à intervalle. Or selon ces mêmes auteurs :

« Les données issues d'une échelle de Likert [...] sont analysées [...] en calculant un score composite (somme ou moyenne) à partir de 4 items de type Likert ou plus » (ibid., p. 3, traduction libre).

Dans la présente étude, les échelles de type Likert que nous utilisons – c'est-à-dire l'échelle SEPE et les sous-échelles SEP(scol), SEP(srl) et SEP(coop) – comportent respectivement 7, 10 et 5 items de type Likert. Selon ces auteurs, la moyenne peut donc être utilisée ici comme indicateur de tendance centrale.

4.2.7.5 Compromis adopté dans la présente recherche

Compte tenu de l'absence de consensus en la matière et de la relative pertinence des deux indicateurs de tendance centrale avec ce type de données, nous avons choisi de pratiquer nos analyses inférentielles à partir des moyennes, puis à partir des médianes.

Concrètement, lorsque nous sommes amenés à traiter statistiquement, non pas des valeurs brutes (*e.g.* scores de scripteurs, données ordinales issues de l'échelle SEC-IAP à item unique), mais des indicateurs de tendance centrale – *e.g.* lorsque pour un échantillon donné il s'agit de comparer en intragroupe le score individuel obtenu à *n* temps de mesure aux 37 items de l'échelle SEPE – nous comparons tous les scores individuels médians de l'échantillon concerné, tout en complétant cette analyse avec celle du score individuel moyen.

4.2.8 Seuil de significativité retenu pour la valeur de p

Le seuil de significativité retenu pour la valeur de p est de .05 – sans toutefois considérer que cette valeur-seuil (« *cut-off point* ») constitue une référence absolue, ni une preuve catégorique pour rejeter l’hypothèse nulle et prouver l’hypothèse alternative. Il s’agit plutôt d’envisager la valeur de p dans le cadre d’un continuum (O’Brien *et al.*, 2015), comme illustré dans le Tableau 5.

$P \geq 0.1$	Absence of evidence against the null hypothesis: data consistent with the null hypothesis
$0.05 \leq P < 0.1$	Low evidence against the null hypothesis in favour of the alternative
$0.01 \leq P < 0.05$	Moderate evidence against the null hypothesis in favour of the alternative
$0.001 \leq P < 0.01$	Strong evidence against the null hypothesis in favour of the alternative
$P < 0.001$	Very strong evidence against the null hypothesis in favour of the alternative

Tableau 5 – Grille indicatrice pour l’interprétation de la valeur de p , d’après (Singh, 2013)

4.2.9 Approche pour mesurer la cohérence interne de l’échelle SEPE

4.2.9.1 Problème posé par l’alpha de Cronbach ici

Bien que l’alpha de Cronbach de (Cronbach, 1951) soit couramment utilisé en sciences humaines pour évaluer la cohérence interne d’une échelle ; il est généralement considéré comme plus légitime lorsque les données suivent une distribution normale (Hayes & Coutts, 2020). Pour déterminer si des données sont normalement distribuées, on peut utiliser le test de Shapiro-Wilk (Shapiro & Wilk, 1965) ; toutefois nous avons vu *supra* que la pertinence de ce test est discutable compte tenu de la (très) petite taille de nos échantillons. En tout état de cause, le test de Shapiro-Wilk que nous avons effectué³⁶ (Tableau 6) révèle une distribution non-normale de nos données SEPE ($N = 282$, $W = 0.80$, $p < .001$). Ainsi, l’utilisation du seul coefficient de Cronbach peut poser problème ici.

N	282
Manquants	13
Taux de déperdition	0.001

³⁶ Ce résultat est présenté ici, dans la partie « Choix méthodologiques », dans la mesure où il ne s’agit pas d’un résultat directement lié à nos prédictions mais d’un préalable qui détermine la méthode statistique employée pour mesurer la cohérence interne.

W de Shapiro-Wilk	0.80
Valeur <i>p</i> de Shapiro-Wilk	< .001

Tableau 6 – Résultats du test de Shapiro-Wilk pour l'échelle SEPE

4.2.9.1 Combinaison avec le coefficient oméga de McDonald

L'oméga de McDonald (McDonald, 2013, 2014) constitue une alternative souvent considérée comme plus robuste que l'alpha de Cronbach (Dunn *et al.*, 2014), en particulier avec des échelles Likert comportant plus de 4 items (Xiao & Hau, 2023). Pour renforcer la validité de nos tests de cohérence interne, nous avons donc utilisé ces deux coefficients en vérifiant s'ils convergent.

4.2.9.2 Population concernée par les analyses de cohérence interne

Les deux analyses de fiabilité précitées sont pratiquées sur l'ensemble des 282 jeux de réponses collectés avec l'échelle SEPE. Cette population (N = 282) correspond en effet à l'ensemble des élèves ayant répondu au questionnaire SEPE, tous échantillons confondus, pour les quatre temps de mesure T1, T2, T3 et T4³⁷. Nous appliquons ensuite la même procédure pour les trois sous-échelles afférentes qui revêtent un intérêt spécifique pour notre étude : c'est-à-dire les sous-échelles SEP(scol), SEP(srl) et SEP(coop).

4.2.10 Approche statistique pour tester l'hypothèse HO 1

Pour tester l'hypothèse HO 1 – c'est-à-dire « La coévaluation numérique favorise l'amélioration des compétences en français » –, nous utilisons deux indicateurs, à savoir :

- Le score individuel de scripteur³⁸ ;
- Le SEP(scol)³⁹.

³⁷ La population concernée par ces analyses de cohérence interne, portant sur les scores SEPE du super-échantillon (SE) tous temps de mesure confondus (*i.e.* N = 282), est à distinguer de la population moins nombreuse des échantillons appariés à partir desquels les analyses inférentielles sont ensuite pratiquées ($4 \leq N \leq 62$).

³⁸ Pour rappel, nous appelons « score de scripteur » le score obtenu par un élève participant pour sa production écrite. Dans le présent protocole, comme expliqué *supra*, c'est l'hétéro-évaluation de l'enseignant qui détermine ce score.

³⁹ SEP(scol) : Sentiment d'efficacité personnelle pour la réussite scolaire

4.2.10.1 Score individuel de scripteur

Eu égard au SEP(scol), nous recueillons des données numériques que nous analysons quantitativement. En l'occurrence, nous comparons les scores individuels de scripteurs des élèves participants :

- En intragroupe :
 - pour chaque échantillon (A, A', B et C) ;
 - pour le super-échantillon (SE) ;
- Sur valeurs brutes (scores exprimés en pourcentages) ;
- Avec le test de Wilcoxon (comparaison pair-à-pair entre deux temps de mesure), c'est-à-dire :
 - T2 (1^{ère} production écrite) vs. T4 (2^{ème} production écrite) ;
 - T4 (2^{ème} production écrite) vs. T6 (3^{ème} production écrite)⁴⁰ ;
- Avec le test de Friedman (approche longitudinale), c'est-à-dire :
 - T2, T4 et T6⁴¹.

4.2.10.2 SEP(scol)

Nous recueillons par ailleurs des données textuelles autorapportées, en l'occurrence les réponses des élèves participants à la sous-échelle SEP(scol). Comme précisé *supra*, ces données textuelles issues d'une échelle de Likert à cinq points sont transposées sous forme de données numériques ordinales et analysées quantitativement en tant que données à intervalle :

- En intragroupe :
 - pour chaque échantillon (A, A', B et C) ;
 - pour le super-échantillon (SE) ;
- Sur scores individuels :
 - médians ;
 - moyens ;
- Avec le test de Wilcoxon (comparaison pair-à-pair entre deux temps de mesure), c'est-à-dire :
 - T1 vs. T2
 - T3 vs. T4

⁴⁰ Echantillon C uniquement.

⁴¹ Echantillon C uniquement.

- T5 vs. T6⁴²
- Avec le test de Friedman (approche longitudinale), c'est-à-dire :
 - T1, T2, T3 et T4
 - T1, T2, T3, T4, T5 et T6⁴³

4.2.11 Approche statistique pour tester l'hypothèse HO 2

Nous testons l'hypothèse opérationnelle HO 2 – c'est-à-dire « La coévaluation numérique améliore l'autodirection des apprentissages » – en la déclinant en trois prédictions, c'est-à-dire « La coévaluation numérique améliore... » :

- « ...l'autorégulation des apprentissages » (HO 2a) ;
- « ...la métacognition » (HO 2b) ;
- « ...le sentiment d'efficacité personnelle » (HO 2c).

Pour tester HO 2a, HO 2b et HO 2c, nous recueillons des données textuelles et/ou numériques, traitées quantitativement et/ou qualitativement, selon l'approche statistique décrite dans les sections qui suivent. Les résultats significatifs obtenus relativement à chacune de ces trois hypothèses opérationnelles sont ensuite combinés afin de tester l'hypothèse HO 2.

4.2.11.1 Approche statistique pour tester l'hypothèse HO 2a

Pour tester l'hypothèse HO 2a – c'est-à-dire « La coévaluation numérique améliore l'autorégulation des apprentissages », nous recueillons les données textuelles issues de la sous-échelle SEP(srl) du questionnaire SEPE⁴⁴, mesurant le sentiment d'efficacité personnelle pour l'autorégulation des apprentissages via une échelle de Likert à cinq points. Nous transposons ces données textuelles sous la forme de données ordinales, que nous traitons quantitativement en tant que données à intervalle. Pour comparer les scores obtenus, nous procédons comme suit :

- Nous utilisons :
 - le test de Wilcoxon pour comparer de deux échantillons appariés ;
 - le test de Friedman pour comparer trois échantillons appariés ou plus ;
- Ces tests sont pratiqués :

⁴² Echantillon C uniquement.

⁴³ Echantillon C uniquement.

⁴⁴ SEPE : Sentiment d'efficacité personnelle des enfants, pré-adolescents et adolescents.

- sur scores individuels médians :
- sur scores individuels moyens ;
- en intragroupe, pour chaque échantillon, c'est-à-dire A, A', B, C et SE ;
- à chaque temps de mesure, c'est-à-dire à T1, T2, T3, T4, T5 et T6.

4.2.11.2 Approche statistique pour tester l'hypothèse HO 2b

Nous testons l'hypothèse HO 2b – c'est-à-dire « La coévaluation numérique améliore la métacognition » – à partir de plusieurs indicateurs, *e.g.* le score de coévaluateur, le score d'auto-évaluateur et le score de scripteur. Toutes les données afférentes sont numériques, exprimées en valeurs brutes. Elles sont analysées quantitativement, au moyen des tests de Wilcoxon, Friedman et Spearman, selon l'approche détaillée ci-après.

Pour commencer nous nous basons sur les scores obtenus par les élèves participants en tant que coévaluateur d'une part, et en tant qu'auto-évaluateur d'autre part. En effet, comme nous l'avons vu *supra*, la capacité des élèves à évaluer leurs pairs ou à s'évaluer eux-mêmes avec justesse constituent des indicateurs fiables pour nous renseigner sur leurs métacompétences. Or le dispositif expérimental de coévaluation numérique permet de comparer finement les hétéro-, auto- et inter-évaluations pour déterminer un score ad hoc, à savoir le score de coévaluateur.

Pour mesurer la capacité des élèves participants à évaluer autrui avec justesse, nous comparons donc en intragroupe les scores de coévaluateurs :

- Des échantillons A, A', B, C et du super-échantillon (SE) appariés ;
- En pratiquant :
 - le test de Wilcoxon, pour comparer :
 - T2 vs. T4 ;
 - T4 vs. T6⁴⁵ ;
 - le test de Friedman, pour comparer :
 - T2, T4 et T6⁴⁶.

Pour rappel, comme indiqué *supra*, l'auto-évaluation a un caractère optionnel dans le protocole co-construit avec les enseignantes partenaires. Cette option n'est utilisée qu'à trois reprises :

- avec l'échantillon A' à T4 ;

⁴⁵ Echantillon C uniquement.

⁴⁶ Echantillon C uniquement.

- avec l'échantillon B à T2 et à T4.

Les scores d'auto-évaluateurs ne sont utiles que s'ils sont comparés aux scores de scripteurs, c'est-à-dire lorsqu'on mesure la capacité des élèves à s'évaluer eux-mêmes avec justesse – l'hétéro-évaluation de l'enseignante partenaire faisant office d'étalon de mesure.

Au moyen du test de Spearman, nous cherchons donc à établir une corrélation entre les scores de scripteurs et les scores d'auto-évaluateurs :

- A T2, pour l'échantillon B ;
- A T4, pour les échantillons A', B et pour l'échantillon composite afférent (A' + B).

De même, nous pratiquons le test de Spearman sur les scores de scripteurs et les scores de coévaluateurs, afin d'apprécier dans quelle mesure les compétences des élèves en production écrite sont corrélées aux métacompétences cibles :

- A T2 et à T4 ;
- Sur les échantillons appariés A, A', B, C et sur le super-échantillon (SE).

4.2.11.3 Approche statistique pour tester l'hypothèse HO 2c

Pour tester l'hypothèse HO 2c – c'est-à-dire « La coévaluation numérique améliore le sentiment d'efficacité personnelle » –, nous recueillons les réponses des élèves participants aux 37 items du questionnaire SEPE⁴⁷. Ces données textuelles sont collectées via une échelle de Likert à cinq points, puis transposées sous la forme de données ordinales. Nous les analysons ensuite quantitativement en tant que données à intervalle, au moyen des tests de Wilcoxon et de Friedman, pratiqués sur scores individuels médians et sur scores individuels moyens, afin d'étudier l'évolution des scores à T1, T2, T3 et T4 – ainsi qu'à T5 et T6 pour l'échantillon C. Comme indiqué *supra*, les temps de mesure précités correspondent aux passations pré/post du questionnaire auprès des élèves participants, avant et après chaque itération expérimentale : les chiffres impairs renvoient à une pré-passation ; les chiffres pairs à une post-passation.

4.2.12 Approche statistique pour tester l'hypothèse HO 3

Pour tester l'hypothèse HO 3 – c'est-à-dire « La coévaluation numérique favorise les apprentissages coopératifs » –, nous disposons de deux indicateurs autorapportés :

⁴⁷ SEPE : Sentiment d'efficacité personnelle des enfants, pré-adolescents et adolescents.

- le score individuel à la sous-échelle SEP(coop)⁴⁸ ;
- le score individuel à l'échelle unidimensionnelle à item unique mesurant le sentiment d'efficacité collective selon l'indice agrégatif personnel (SEC-IAP).

Initialement, ces deux indicateurs sont mesurés par l'entremise des données textuelles – c'est-à-dire les réponses aux différents items – collectées via une échelle de Likert à cinq points. Comme précédemment, ces données textuelles sont transposées sous forme de données ordinales puis analysées quantitativement en tant que données à intervalle.

Pour chaque échantillon et pour le super-échantillon (SE) appariés, afin de comparer en intragroupe les valeurs d'un temps de mesure à l'autre, nous utilisons :

- le test de Wilcoxon (pour deux temps de mesure comparés pair-à-pair), avec le coefficient de corrélation entre rangs bisériés ($-1 < r < 1$) comme estimateur de taille d'effet (si une différence significative est détectée), c'est-à-dire pour nous renseigner sur l'ampleur de la hausse ou de la baisse constatée ;
- le test de Friedman (pour trois temps de mesure ou plus), avec le test de Durbin-Conover en analyse post-hoc, lorsqu'une différence significative est détectée, dans le but d'affiner l'analyse avec des comparaisons pair-à-pair ;

Pour rechercher une corrélation entre les deux indicateurs précités, nous utilisons le *rho* de Spearman ($-1 < r < 1$).

En ce qui concerne les scores à la sous-échelle SEP(coop), nous avons pratiqué ces tests sur scores individuels médians, puis sur scores individuels moyens. Cette distinction ne s'appliquait pas pour les scores à l'échelle SEC-IAP, dans la mesure où il s'agissait d'une échelle à item unique générant une valeur brute.

4.2.13 Approche statistique pour tester l'hypothèse HO 4

Pour tester l'hypothèse 4 – c'est-à-dire « La coévaluation numérique génère des conflits instrumentaux et permet de les réguler » –, nous recueillons plusieurs types de données textuelles, auprès des élèves participants d'une part, et auprès des enseignants partenaires d'autre part. Compte tenu de la variété des données utilisées pour tester l'hypothèse HO 4 – à la fois en termes de publics cibles, de modalités d'enquête et de méthodes d'analyse –, ces données n'ont pas fait l'objet d'un traitement statistique systématisé. Lorsqu'elles sont issues de questions fermées et collectées via une échelle de Likert, les données textuelles sont

⁴⁸ SEP(coop) : Sentiment d'efficacité personnelle pour les apprentissages coopératifs.

transposées sous forme de données ordinales puis analysées quantitativement en tant que données à intervalle. Lorsque les données proviennent de questions ouvertes, d'entretiens semidirigés ou de focus groupes, elles sont traitées qualitativement : par analyse catégorielle de contenus, par analyse textométrique ou à la manière d'un « récit de vie »⁴⁹, selon ce qui semble le plus adapté à la nature des données.

4.3 Focus groupe (échantillon C')

Entre T4 et T5, un focus groupe est mis en œuvre avec la 6^e coopérative du collège C. Cette autre classe de l'enseignante C, ci-après dénommée « échantillon C' » n'expérimente pas la coévaluation numérique pendant la période expérimentale (année N), mais est susceptible de l'expérimenter en année N+1. Après que le dispositif de coévaluation numérique leur a été brièvement présenté par le chercheur-intervenant, le recueil des avis de ces élèves est donc jugé pertinent, à fortiori dans la mesure où ils en ont déjà entendu parler via leurs pairs, élèves de 5^e coopérative – c'est-à-dire via l'échantillon C, qui prend part au présent protocole de recherche. Les données recueillies dans le cadre de ce focus groupe font l'objet d'une prise de note de la part du chercheur-intervenant, en vue d'une brève analyse thématique.

4.4 Retours écrits sur expérience (échantillon C)

Avec l'échantillon C (N = 23), à l'issue des trois itérations du protocole semi-expérimental menées pendant l'année N, les élèves s'installent en salle informatique, un demi-groupe après l'autre (N = 12, puis N = 11), pendant 55 minutes chacun.

Pour rappel, l'échantillon C est le seul avec lequel nous avons pu mettre en œuvre le protocole de recherche à trois reprises (vs. deux itérations avec A, A' et B). Les élèves de cette classe de 5^e sont donc ceux qui ont le plus d'expérience avec le dispositif expérimental de coévaluation numérique, c'est-à-dire ceux dont le retour sur expérience est susceptible d'être le plus riche. En outre, le chef de l'établissement C a consenti à un aménagement d'emploi du temps, de sorte que ces deux séances supplémentaires de 55 minutes en demi-groupe puissent

⁴⁹ Les analyses quantitative, catégorielle et textométrique constituent effectivement des modalités de traitement des données, tandis que le récit de vie renvoie à priori à un type de matériau ou à un format d'enquête. La justification de cette apparente disparité est fournie infra, à la fin de ce chapitre portant sur le « Recueil et traitement des données », dans la section dédiée aux « Entretiens semidirigés et focus groupe avec les enseignantes partenaires ».

avoir lieu, en l'absence de l'enseignante C. Ce fait constitue une exception dans la présente recherche.

Pendant une trentaine de minutes – après avoir consulté leur rapport d'évaluation individuel dans Moodle Atelier (notes de scripteur, notes de coévaluateur, formulaires d'évaluation et écrits commentatifs) et l'avoir enregistré dans leur dossier individuel sur l'intranet du collège –, les élèves sont invités dans un premier temps à écrire à leur enseignante un courriel dans lequel ils rendent compte de leur expérience et de leur ressenti avec ce dispositif. Deux entrées thématiques délibérément simples sont suggérées dans la consigne d'écriture de ces retours sur expérience. L'objectif est de permettre aux élèves participants d'utiliser leurs propres mots pour éclairer la genèse instrumentale du dispositif expérimental en termes de conflictualité, mais aussi en termes de régulations – étant entendu que les élèves eux-mêmes n'ont pas conscience du cadre théorique auquel on se réfère, c'est-à-dire l'approche instrumentale en EIAH. Ces deux entrées thématiques sont les suivantes :

- Ce que j'ai apprécié...
- Ce que je n'ai pas apprécié...

Selon cette approche binaire, les données textuelles recueillies génèrent deux types de retours sur expérience : l'un mélioratif et l'autre péjoratif. Des verbatim issus de ces retours sur expérience sont consultables en Annexes 4. L'orthographe n'est pas rectifiée afin de préserver la forme originelle et l'authenticité de ces écrits, et dans la mesure où les erreurs orthographiques ou syntaxiques qui s'y trouvent n'en altèrent pas significativement l'intelligibilité.

En outre, s'agissant de courriels adressés à leur enseignante, cette dernière a pris l'initiative de leur répondre individuellement. Ces réponses de l'enseignante C sont susceptibles de fournir un éclairage supplémentaire eu égard à la dynamique collaborative de la présente recherche ; elles sont donc incluses dans les verbatim.

Au moyen du logiciel IRaMuTeQ, les retours sur expérience des élèves font *in fine* l'objet d'une analyse textométrique, c'est-à-dire en appliquant une approche quantitative à ces données textuelles, au moyen du traitement automatique des langues (TAL). Pour ce faire, les écrits d'élèves sont préalablement formatés, indexés et codés, comme précisé dans le chapitre Apparatus. Les erreurs orthographiques et syntaxiques y sont également rectifiées à cet effet.

4.5 Focus groupe (échantillon C)

Pendant une vingtaine de minutes, dans le prolongement des préparatifs décrits *supra*, un focus groupe est mis en œuvre à T6 avec chaque demi-groupe de la classe C (élèves de 5^e), à partir des deux entrées thématiques suggérées pour les retours écrits sur expérience. Dans la mesure où ces derniers constituaient en quelque sorte un écrit de travail déclenchant le brainstorming, ces entrées thématiques simples servent cette fois de point de départ aux échanges, afin d'éclairer la genèse instrumentale du dispositif expérimental, c'est-à-dire pour mettre au jour les mécanismes de conflits et de régulations instrumentaux afférents. Sachant que l'enseignante n'est pas présente lors de ces échanges, il s'agit d'encourager une certaine liberté de parole, pour qu'au-delà des conflits instrumentaux rencontrés et des régulations mises en œuvre en classe, on puisse favoriser l'émergence de propositions alternatives de régulation instrumentale, *e.g.* fondées sur de nouvelles catachrèses issues de l'intelligence collective des élèves eux-mêmes (Bellet, 2022).

Les deux focus groupes font l'objet d'un enregistrement audio intégral ; cependant le fichier audio n'est pas versé en annexe pour préserver l'anonymat des locuteurs. Seul des transcripts anonymisés figurent en Annexes 5. Ces transcripts ont été réalisés manuellement et leur intégrité discursive a été préservée, pour laisser entendre la voix authentique des élèves, avec leurs mots et expressions.

Traitées séparément au moyen d'analyses catégorielles de contenus, les données textuelles ainsi recueillies sont ensuite regroupées dans la mesure où il s'agit de deux sous-groupes appartenant au même groupe classe, ayant expérimenté le dispositif de coévaluation numérique en classe entière, dans des conditions similaires, à trois reprises.

4.6 Entretiens semidirigés et focus groupe (enseignantes partenaires)

Nous menons à la suite de chaque expérimentation en classe des entretiens semidirigés, en présentiel ou en distanciel, avec les enseignantes partenaires. Ces entretiens ont pour but de recueillir les réactions des enseignantes partenaires « à chaud », alors que les premiers ressentis relatifs au vécu expérimental sont encore vivaces. En outre, à la fin de la période expérimentale (année N) un focus groupe est organisé en visioconférence pour permettre aux trois enseignantes partenaires et au chercheur-intervenant d'échanger.

À chaque fois, une grille d'analyse commune est utilisée, axée sur les trois composantes de la genèse instrumentale du dispositif de coévaluation que nous avons expérimenté, c'est-à-dire : artéfact didactique, artéfact pédagogique et artéfact technique – ce dernier faisant l'objet d'un questionnement spécifique eu égard aux conflits et régulations instrumentaux.

La méthode d'analyse retenue pour ces données textuelles ne consiste cependant pas en une approche systématique – *e.g.* une analyse textométrique ou une analyse catégorielle de contenus. Au contraire, bien qu'ils soient conçus en amont comme des entretiens semidirigés et un focus groupe, le caractère paritaire, horizontal de la relation entre le chercheur-intervenant et les enseignantes partenaires confère à ces modalités de recueil un caractère très libre, qui revient de facto à « raconter à une autre personne un épisode de son expérience vécue » (Burrick, 2010, p. 9).

Les enregistrements audios de ces échanges (Annexes 6 et 7) servent ainsi de base à une analyse de type « récit de vie ». Souvent utilisée en recherche en sciences de l'éducation et de la formation, notamment pour revenir sur un épisode vécu en formation par des personnels enseignants (Burrick, 2010), cette approche semble pertinente dans le cadre de cette recherche-intervention collaborative (Bertaux, 2003), dans la mesure où une dynamique de recherche-action étroitement concertée est mise en œuvre avant, pendant et après la phase expérimentale.

L'analyse qualitative à la manière d'un récit de vie paraît ainsi la plus à même de rendre compte de la richesse et de la singularité de chaque témoignage. Pour ce faire, nous cherchons à rendre compte du discours des enseignantes partenaires avec la plus grande fidélité possible, notamment en ce qui concerne l'ordre naturel d'apparition des idées, « au fil de l'eau ». La structure initialement prévue pour les entretiens semidirigés et le focus groupe sert néanmoins de fil conducteur aux échanges, ce qui procure aux récits de vie afférents une forte cohérence au plan thématique.

Tout en gardant à l'esprit l'écart qui peut exister entre pratiques déclarées et pratiques effectives, il est possible que cette dimension confère aux entretiens semidirigés et au focus groupe une valeur spécifique au plan épistémologique en général, et praxéologique en particulier. Notamment, dans la mesure où les enseignantes partenaires sont pleinement associées à la recherche, en tant qu'acteurs à part entière, l'approche instrumentale qui sert de cadre à ces échanges est abordée en toute transparence et constitue une base commune de travail et de réflexion tout au long de la présente recherche.

Les conditions de ces entretiens semidirigés et de ce focus groupe sont résumées dans le Tableau 7.

Modalité de recueil	Temps de mesure	Participants	Durée	Lieu
Entretien semidirigé	T2 (année N)	Enseignante A Chercheur-intervenant	5 min	Collège A (salle informatique)
Entretien semidirigé	T2 (année N)	Enseignante C Chercheur-intervenant	70 min	Collège C (salle informatique)
Entretien semidirigé	T4 (année N)	Enseignante B Chercheur-intervenant	50 min	Distanciel (téléphone)
Entretien semidirigé	T4 (année N)	Professeure-documentaliste (collège A) Chercheur-intervenant	30 min	Centre de documentation et d'information (CDI) du collège A
Entretien semidirigé	T6 (année N)	Enseignante C Chercheur-intervenant	90 min	Collège C (salle informatique)
Focus groupe	T6 (année N)	Enseignante A Enseignante B Enseignante C Chercheur-intervenant	80 min	Distanciel (visioconférence)
Entretien semidirigé	Année N+1	Enseignante B Chercheur-intervenant	60 min	Distanciel (téléphone)

Tableau 7 – Conditions de déroulement des entretiens semidirigés et du focus groupe avec les enseignantes partenaires

5 Apparatus

5.1 Moodle

La coévaluation numérique individualisée a été pratiquée à l'aide de Moodle, version 3.11, une plateforme d'apprentissage en ligne (« *learning management system* »). Plus spécifiquement, la coévaluation était pratiquée au moyen de Moodle Atelier (« *Moodle Workshop* ») un module qui y est intégré, et qui est initialement prévu pour pratiquer l'évaluation entre pairs.

5.1.1 Accès à Moodle Atelier (ENT académique)

Pour les besoins de cette recherche-intervention, suite à la validation du projet par les instances académiques de Montpellier et Paris, les DANE et DRANE de Montpellier et Paris nous ont fourni un accès à leur ENT respectif, via la création d'un compte ad hoc. Nous disposons ainsi des autorisations requises pour créer et éditer des cours sur Moodle, et surtout pour mettre en œuvre des Ateliers Moodle en collaboration avec les enseignantes partenaires.

5.1.2 Catachrèse opérée pour pratiquer la coévaluation avec Moodle Atelier

Nous avons détourné l'usage initial de Moodle Atelier – c'est-à-dire l'inter-évaluation – afin de pratiquer la coévaluation. Nous avons procédé pour ce faire à une catachrèse dont le détail se trouve *supra*, dans notre état de l'art.

5.2 Ordinateurs en salle informatique

L'ensemble du protocole est mis en œuvre dans les salles informatiques⁵⁰ des collèges partenaires en utilisant leurs ordinateurs de type PC et de marque Lenovo, équipés du système d'exploitation Microsoft® Windows™ 7. Tous les ordinateurs disposent d'un accès à l'intranet du collège (pour la sauvegarde des productions écrites, puis leur téléversement dans Moodle Atelier), ainsi que d'un accès à l'ENT académique – où Moodle est installé. Dans l'ensemble, les élèves disposaient de leurs codes d'accès ; lorsque certains élèves ignoraient ou avaient oublié leurs codes, le problème a pu être résolu rapidement dans la plupart des cas.

5.3 Traitement de texte utilisé pour les productions écrites

Le traitement de texte utilisé pour les productions écrites était le logiciel Libre Office. Alternativement, le logiciel Open Office a pu être utilisé par certains élèves, car les deux logiciels étaient parfois installés par défaut sur les ordinateurs des collèges partenaires. Un correcticiel intégré, ainsi que des modules complémentaires dédiés à la correction orthographique et grammaticale (*e.g.* le logiciel libre [Grammalecte](#)) peuvent être mobilisés dans ces traitements de texte, toutefois cette dimension n'a pas été prise en compte dans la présente recherche.

5.4 Logiciel utilisé pour le questionnaire en ligne

Pour la passation du questionnaire en ligne, nous avons utilisé le logiciel [Tripetto](#)®. Nous avons effectué ce choix en raison du caractère ergonomique de son interface utilisateur,

⁵⁰ Exceptionnellement, lors d'une panne affectant plusieurs postes informatiques dans l'établissement A, 3 élèves de la classe A' ont été accueillis au Centre de documentation et d'information (CDI) du collège. Ces circonstances font l'objet d'une analyse ad hoc, *infra*.

avec notamment, pour le public de collégiens que nous ciblons, des fonctionnalités relevant de la ludification, *e.g.* le déroulement du questionnaire sous la forme d'un dialogue, avec des bulles évoquant l'univers de la bande-dessinée et un avatar représentant le chercheur posant ses questions (Figure 42).

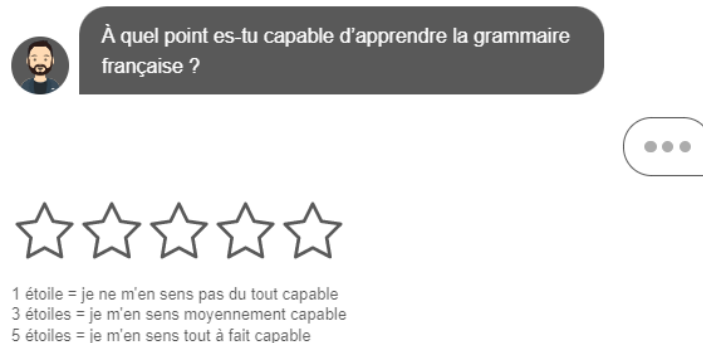


Figure 42 – Interface utilisateur dans Tripetto® : éléments de ludification

Le questionnaire a ainsi été administré en présentiel, en salle informatique, sur les postes informatiques du collège. Un lien URL était cliquable à même l'interface de Moodle, immédiatement avant et immédiatement après l'activité de coévaluation (Figure 43).

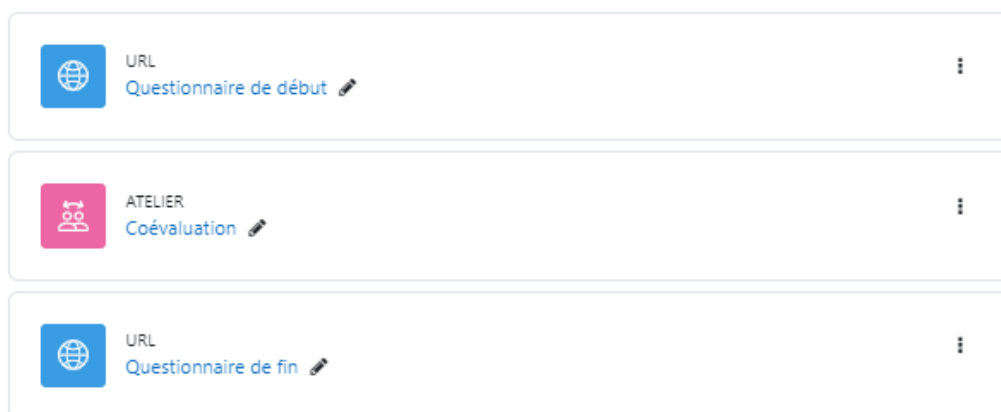


Figure 43 – Accessibilité du questionnaire dans Moodle

Pour éviter des problèmes de surcharge, observés lorsque toute une classe se connecte en même temps au même questionnaire en ligne via un lien unique, nous divisons la classe en deux, suivant l'ordre alphabétique. En créant un doublon du questionnaire, nous pouvons ainsi envoyer les élèves vers l'original ou le doublon, en fonction de l'initiale de leur nom de famille (Figure 44), ce qui divise le trafic par deux.



Figure 44 – Dédoublage du questionnaire pour pallier les surcharges sur le serveur du collège

5.5 Logiciel utilisé pour les analyses statistiques

Pour les analyses statistiques, nous utilisons le logiciel *open-source* [Jamovi](#) dans sa version « *desktop* » 2.3.21. Ce logiciel est basé sur le langage R. Grâce à une interface graphique ergonomique, il rend le calcul de statistiques inférentielles plus accessible aux non-spécialistes.

5.6 Logiciel utilisé pour les analyses textométriques

Pour les analyses textométriques pratiquées sur les retours écrits sur expérience de l'échantillon C à T6, nous utilisons le logiciel libre de textométrie IRaMuTeQ (Interface de R pour les analyses Multidimensionnelles de Textes et de Questionnaires), développé par (Ratinaud, 2009), dans sa version 0.7 alpha 2 (2020) pour Windows. Le logiciel étant basé sur le langage statistique R, il nécessite l'installation préalable de R 4.1.3 (<http://www.r-project.org>). L'utilisation du logiciel implique de se familiariser avec certains concepts propres au traitement automatique des langues (TAL), d'où l'utilisation de guillemets, *infra*, pour désigner certains termes techniques propres au fonctionnement du logiciel – *e.g.* « thématique » ou « segmentation ».

5.6.1 Préparation et codage du corpus

Pour les besoins de l'analyse qualitative (analyse catégorielle de contenus), les erreurs orthographiques des retours écrits sur expérience sont rectifiées et les textes sont codés et classés pour mettre en évidence la polarisation des contenus : Retours positifs vs. Retours

négatifs. Pour ce faire, par l'entremise du traitement de texte Libre Office, nous formatons les retours écrits des élèves au format .txt, avec l'encodage universel UTF-8 et en utilisant un codage propre à IRaMuTeQ, *e.g.* des séparateurs de textes (****). On ajoute des métadonnées, en spécifiant des « thématiques ». Ce formatage a pour objectif d'augmenter la granularité de l'analyse en créant *in fine* un sous-corpus pour chacun des deux types de retours sur expérience recueillis.

5.6.2 Non-segmentation du corpus

Les textes du corpus peuvent faire l'objet d'un processus de « segmentation » dans le logiciel ; cependant compte tenu du fait qu'il s'agit de retours sur expérience courts, en provenance d'élèves de 5^e, formulés au moyen de phrases brèves, nous n'avons pas créé de segments de textes mais opté pour une conservation de l'unité textuelle d'origine : chaque retour d'élève conserve ainsi son intégrité phrastique lors de l'analyse textométrique.

5.6.3 Indexation du corpus

Le corpus est ensuite « indexé » dans IRaMuTeQ. Le processus d'indexation implique notamment une transformation du corpus en caractères minuscules et une suppression des apostrophes. Ce processus de « lemmatisation » permet de réduire les diverses formes à un lemme commun, neutre (*e.g.* infinitif pour les verbes conjugués, ou masculin/singulier pour les substantifs) ; ceci afin de réduire la variabilité du langage et d'en faciliter le traitement automatisé. Le logiciel s'appuie sur des dictionnaires pour opérer cette lemmatisation, ce qui permet de distinguer les formes dites « pleines » (*e.g.* verbes, adjectifs, substantifs...) – appelées « formes actives » dans le logiciel –, des « mots-outils » (*e.g.* « et », « avec », « à »...), appelés « formes supplémentaires » dans IRaMuTeQ.

5.6.4 Sous-corpus thématiques

Grâce au formatage et au codage d'un corpus, il est possible de le subdiviser dans IRaMuTeQ en différents sous-corpus, en fonction de variables (*e.g.* sexe : h/f) ou de thématiques (*e.g.* retour mélioratif, retour péjoratif), selon l'objet de la recherche. Comme indiqué *supra*, c'est ici la seconde approche qui est utilisée. En effet, le corpus se prêtait moins à une analyse par variables, en raison de sa nature, de sa taille et de l'objectif de recherche afférent.

5.6.5 Utilisation des tirets bas

Le codage du corpus en amont offre la possibilité de substituer des tirets bas aux espaces, entre certains vocables choisis, pour conserver leur cotexte – c’est-à-dire leur environnement syntaxique immédiat – *e.g.* « découvrir_les_textes_des_camarades ». Ici, nous avons parfois recours à ces tirets bas dans la mesure où nous ne visons pas seulement les lexèmes considérés pour eux-mêmes, mais également « le sémantisme des unités lexicales prises en discours » (Dias Loguercio, 2012 : 20).

5.6.6 Nuages de mots

Des nuages de mots peuvent être générés. Cette approche exploratoire se limite à représenter graphiquement la fréquence relative des mots du corpus. Nous générons ici trois nuages de mots : un pour le corpus global et un pour chaque sous-corpus thématique. Nous obtenons ainsi un aperçu de la nature du corpus, qui nécessite cependant d’être approfondi.

5.6.7 Analyse de similitude

Dans IRaMuTeQ, il est possible d’opérer une analyse de similitude (ADS), dont l’objectif est « d’étudier la proximité et les relations entre les éléments d’un ensemble, sous forme d’arbres » (Marchand & Ratinaud, 2012, citant Degenne *et al.*, 1973; Flament & Rouquette, 2003). Nous réalisons ici une ADS pour prolonger l’analyse exploratoire réalisée avec les trois nuages de mots.

5.7 Ordinateur utilisé pour la présente recherche

Les analyses statistiques et textométriques sont réalisées sur un PC ASUS® VivoBook™ équipé d’un processeur 11^e génération Intel® Core™ i5-1135G7, disposant d’une fréquence d’horloge de 2,40 GHz et de 8 Go de RAM. Le PC est équipé du système d’exploitation Microsoft® Windows™ 11 (64 bits).

5.8 Entrepôt de données

L’entrepôt de données utilisé pour le dépôt des enregistrements audios (entretiens semidirigés et focus groupe avec les enseignantes partenaires) est l’entrepôt disciplinaire [Nakala](#), de la plateforme [Huma-Num](#).

6 Aspects déontologiques

Conformément au Règlement général sur la protection des données (RGPD), aux recommandations de la CNIL et aux préconisations de l'école doctorale 58 de l'université Paul-Valéry Montpellier 3 en la matière, nous faisons remplir et signer à chaque enquêté un formulaire de consentement libre et éclairé. Pour garantir la répliquabilité de notre recherche, il y est précisé que les données collectées sont pseudonymisées pendant la durée de la recherche et assorties d'une durée de conservation de 3 ans à compter de la fin de la collecte des données, afin de permettre l'analyse et la production des résultats de recherche.

Lorsque pour l'analyse approfondie d'un cas particulier, des prénoms d'élèves sont mentionnés, ils ont été changés pour préserver leur anonymat, conformément au formulaire de consentement libre et éclairé signé par leurs parents et conformément au règlement général de protection des données (RGPD).

Pour répondre aux enjeux de la science ouverte et de l'interopérabilité des données, les données collectées peuvent ensuite faire l'objet d'une anonymisation – opération définitive, non réversible, occasionnant la suppression des données identificatoires – pour conservation archivistique sans limite de durée, conformément au code du patrimoine et au RGPD, dans des entrepôts de données dédiés à la recherche et disciplinaires, *e.g.* la plateforme Nakala d'Huma-Num pour les sciences humaines et sociales.

A ce titre, les enregistrements audio des entretiens semidirigés et du focus groupe avec les enseignantes partenaires ont été édités pour en retirer tout matériau portant atteinte à l'anonymat des personnes. Par ailleurs, la publication pérenne de ces enregistrements audio dans un entrepôt disciplinaire de données – en l'occurrence l'entrepôt Nakala, de la plateforme Huma-Num – fait l'objet d'une autorisation spécifique dans le formulaire de consentement libre et éclairé signé par les enseignantes partenaires. La licence de publication *Creative Commons* utilisée est « Attribution non commerciale sans dérivatifs 4.0 internationale » ([CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)).

PARTIE IV : RESULTATS

Dans cette partie, nous présentons les résultats de la présente étude hypothèse par hypothèse, à raison d'un chapitre par prédiction. Le Tableau 8 synthétise l'approche par méthodes mixtes qui a permis d'aboutir à l'ensemble de ces résultats, c'est-à-dire :

- en pratiquant des analyses quantitatives sur des données numériques ou textuelles ;
- ou en pratiquant des analyses qualitatives sur des données textuelles.

Comme précisé *supra*, lorsque pour une prédiction donnée nous disposons de plusieurs types de résultats (*e.g.* pour HO 1 ou HO 2), nous procédons à une intégration par combinaison.

Dans chaque chapitre, la présentation des résultats est suivie d'une « discussion locale » – c'est-à-dire circonscrite à l'hypothèse opérationnelle concernée. Ces discussions locales servent de préalable à la Partie V : Discussion générale.

Enfin, les résultats détaillés de la recherche – *e.g.* statistiques descriptives, statistiques inférentielles, analyses post-hoc, matrices de corrélation – sont disponibles en Annexes 1.

Codage	Hypothèses opérationnelles (HO)		Données	Analyses
HO 1	Amélioration des compétences en français		Numériques	Quantitative
			Textuelles	Quantitative
HO 2	Amélioration de l'autodirection des apprentissages			
	HO 2a	Amélioration de l'autorégulation des apprentissages	Textuelles	Quantitative
	HO 2b	Amélioration de la métacognition	Numériques	Quantitative
	HO 2c	Augmentation du sentiment d'efficacité personnelle	Textuelles	Quantitative
HO 3	Amélioration des apprentissages coopératifs		Textuelles	Quantitative
HO 4	Génération et régulation de conflits instrumentaux		Textuelles	Quantitative
			Textuelles	Qualitative

Tableau 8 – Hypothèses opérationnelles, données afférentes et analyses pratiquées

1 Compétences en français (HO 1)

Pour rappel, nous testons l'hypothèse HO 1 – c'est-à-dire « La coévaluation numérique favorise l'amélioration des compétences en français » – au moyen de deux indicateurs distincts.

D'une part, nous recueillons des données numériques que nous analysons quantitativement, en l'occurrence les scores individuels de scripteurs des élèves participants. Eu égard à cet indicateur, en appliquant systématiquement l'approche statistique détaillée *supra*, nous avons pratiqué sept analyses inférentielles, dont deux ont généré un résultat significatif.

Nous recueillons par ailleurs des données textuelles autorapportées, en l'occurrence les réponses des élèves participants à la sous-échelle SEP(scol)⁵¹. En appliquant systématiquement avec cet autre indicateur l'approche statistique détaillée *supra*, nous avons pratiqué 34 analyses inférentielles, dont trois ont généré un résultat significatif.

Ces cinq résultats significatifs sont détaillés dans les sections qui suivent.

1.1 Augmentation modérée du score de scripteur de SE sur T2-T4

Le test de Wilcoxon pratiqué sur échantillons appariés met en évidence une augmentation modérée des scores individuels de scripteurs pour le super-échantillon⁵² (SE) entre T2 et T4 ($p = 0.001$; $r = 0.51$; $N = 58$) – c'est-à-dire entre la première et la deuxième production écrite (Figure 45).

⁵¹ SEP(scol) : Sentiment d'efficacité personnelle pour la réussite scolaire

⁵² Pour rappel, le super-échantillon (SE) est composé de l'ensemble des élèves participants, tous échantillons confondus.

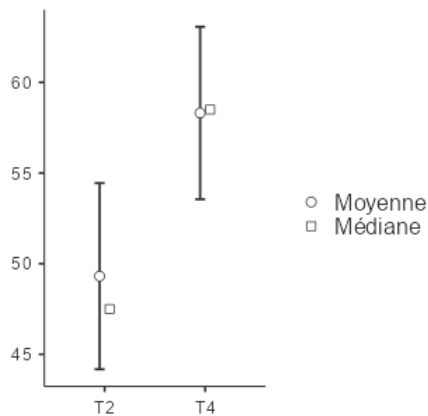


Figure 45 – Augmentation modérée des scores individuels de scripteurs de SE entre T2 et T4

1.2 Différence significative entre les scores de scripteurs de l'échantillon C sur T2-T6

Pour l'échantillon C (N = 22) – avec lequel une troisième itération du protocole a pu être mise en œuvre à T5/T6 –, le test de Friedman indique une différence significative entre les scores individuels de scripteurs des élèves à T2, T4 et T6 ($p = 0.018$). L'analyse post-hoc réalisée au moyen du test de Durbin-Conover montre par ailleurs une significativité accrue de la différence entre les scores individuels de scripteurs à T2 et T6 ($p = 0.005$).

Par ailleurs, lorsqu'on compare les moyennes des scores de scripteurs de l'échantillon C (N = 22) à T2, T4 et T6, on constate que la tendance haussière, précédemment mise en évidence pour le super-échantillon (SE) sur l'intervalle T2-T4, se poursuit au-delà pour l'échantillon C, sur l'intervalle T4-T6 (Figure 46).

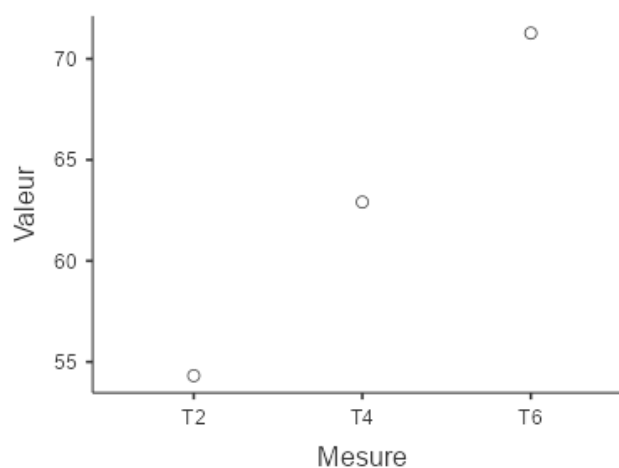


Figure 46 – Augmentation du score de scripteur moyen de C sur T2-T6 (scores moyens exprimés en %, N = 22)

1.3 Augmentation du SEP(scol) de SE sur T3-T4

Le test de Wilcoxon pour échantillons appariés, pratiqué sur les scores individuels moyens obtenus par les élèves du super-échantillon SE (N = 46) à l'échelle SEP(scol), indique une augmentation modérée entre T3 et T4 ($p = 0.001$; $r = 0.663$), comme l'illustre la Figure 47.

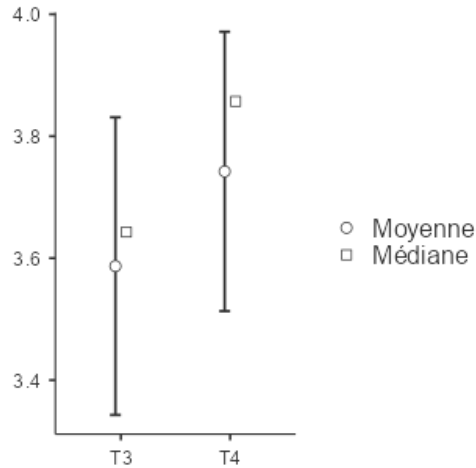


Figure 47 – Augmentation modérée des scores individuels moyens des élèves de SE à la sous-échelle SEP(scol) entre T3 et T4

Ce résultat concernant le super-échantillon (SE) est corroboré par les tests de Wilcoxon pratiqués sur scores individuels moyens pour les élèves des échantillons A' et C sur le même intervalle – c'est-à-dire entre T3 et T4. On constate en effet une forte augmentation du SEP(scol) de A' ($p = 0.032$; $r = 0.822$; N = 12) et une très forte augmentation du SEP(scol) de C ($p = 0.036$; $r = 1$; N = 16). Ces augmentations sont représentées par la Figure 48 et la Figure 49.

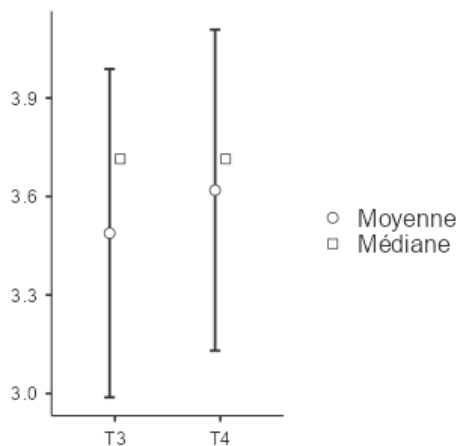


Figure 48 – Forte augmentation des scores individuels moyens des élèves de A' à la sous-échelle SEP(scol) sur T3-T4

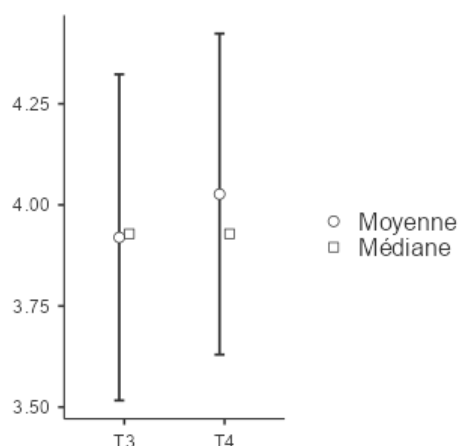


Figure 49 – Très forte augmentation des scores individuels moyens des élèves de C à la sous-échelle SEP(scol) sur T3-T4

1.4 Discussion (HO 1)

Les résultats obtenus sont relativement peu nombreux, notamment en raison de notre approche non-paramétrique et de la perte de puissance statistique qui en résulte. Toutefois, la robustesse statistique des tests utilisés et la valeur-seuil retenue pour p confèrent à ces résultats une certaine solidité. De surcroît, l'approche par méthodes mixtes nous permet ici de combiner divers types de résultats – c'est-à-dire des résultats d'analyses quantitatives pratiquées sur des données textuelles d'une part, et des résultats d'analyses quantitatives pratiquées sur des données numériques d'autre part.

Ainsi, eu égard au sentiment d'efficacité personnelle pour la réussite scolaire, les trois résultats significatifs que nous avons obtenus sont convergents : ils indiquent une augmentation modérée, voire élevée, pour l'ensemble des élèves entre T3 et T4. Or compte tenu du lien entre le SEP(scol) et les compétences réelles des élèves, cette hausse corrobore l'amélioration, établie par ailleurs, des compétences des élèves participants en français – c'est-à-dire l'augmentation des scores individuels de scripteurs de l'ensemble des élèves pendant toute la durée de la période expérimentale.

La combinaison de ces résultats (Tableau 9) nous permet donc de valider l'hypothèse HO 1, c'est-à-dire « La coévaluation numérique favorise l'amélioration des compétences en français ».

Hypothèse	Echant.	Intervalles	Indicateurs	Résultats
HO 1 : Amélioration des compétences en français	SE	T2-T4	Scores de scripteurs	Augmentation modérée
	C	T2-T6	Scores de scripteurs	Différence significative
	A'	T3-T4	SEP(scol)	Forte augmentation
	C	T3-T4	SEP(scol)	Très forte augmentation
	SE	T3-T4	SEP(scol)	Forte augmentation

Tableau 9 – Synthèse des résultats obtenus pour l'hypothèse HO 1 : Amélioration des compétences en français

2 Autodirection des apprentissages (HO 2)

Pour rappel, nous testons l'hypothèse opérationnelle HO 2 – c'est-à-dire « La coévaluation numérique améliore l'autodirection des apprentissages » – en la déclinant en trois prédictions, c'est-à-dire « La coévaluation numérique améliore... » :

- « ...l'autorégulation des apprentissages » (HO 2a) ;
- « ...la métacognition » (HO 2b) ;
- « ...le sentiment d'efficacité personnelle » (HO 2c).

Les résultats significatifs ainsi obtenus sont combinés afin de tester l'hypothèse HO 2.

2.1 Autorégulation des apprentissages (HO 2a)

Pour tester l'hypothèse HO 2a, en appliquant systématiquement l'approche statistique détaillée *supra*, nous avons utilisé l'indicateur SEP(srl)⁵³ pour pratiquer 34 analyses inférentielles, dont quatre ont généré des résultats significatifs.

2.1.1 Augmentation du SEP(srl) de SE sur T1-T2 et T3-T4

Pour les élèves du super-échantillon SE (N = 46), le test de Wilcoxon pour échantillons appariés, pratiqué sur scores individuels médians, révèle notamment une augmentation modérée du SEP(srl) entre T1 et T2 ($p = 0.042$; $r = 0.462$). En outre, pour ce même super-échantillon (SE), le test de Wilcoxon pour échantillons appariés, pratiqué sur scores individuels moyens, indique une légère augmentation du SEP(srl) entre T3 et T4 ($p = 0.042$; $r = 0.384$). La Figure 50 et la Figure 51 illustrent ces résultats.

⁵³ SEP(srl) : Sentiment d'efficacité personnelle pour l'autorégulation des apprentissages

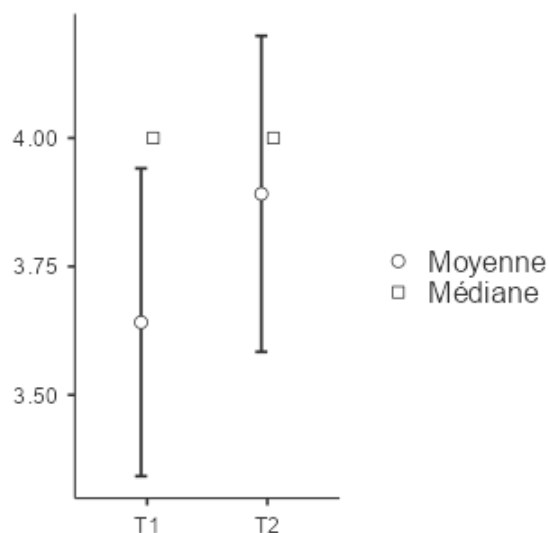


Figure 50 – Augmentation modérée des scores individuels médians des élèves de SE à la sous-échelle SEP(srl) sur T1-T2

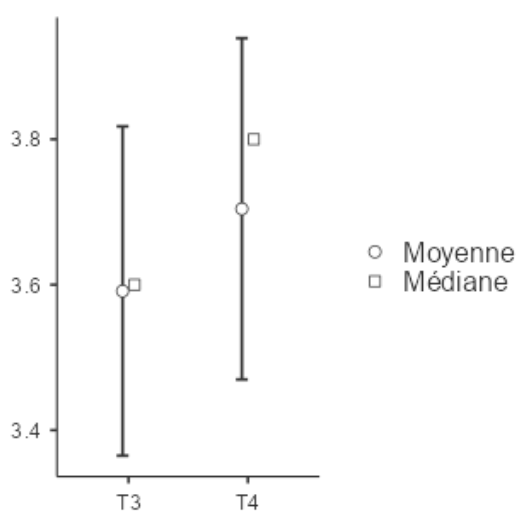


Figure 51 – Légère augmentation des scores individuels médians des élèves de SE à la sous-échelle SEP(srl) sur T3-T4

2.1.2 Forte augmentation du SEP(srl) de B sur T1-T2

Concernant l'échantillon B en particulier ($N = 10$), le test de Wilcoxon pour échantillons appariés, pratiqué sur scores individuels médians, révèle quant à lui une forte augmentation du SEP(srl) entre T1 et T2 ($p = 0.041$; $r = 0.778$). Le test de Wilcoxon pour échantillons appariés, pratiqué sur scores individuels moyens, corrobore ce résultat ($p = 0.037$; $r = 0.745$). La Figure 52 rend compte de ce résultat.

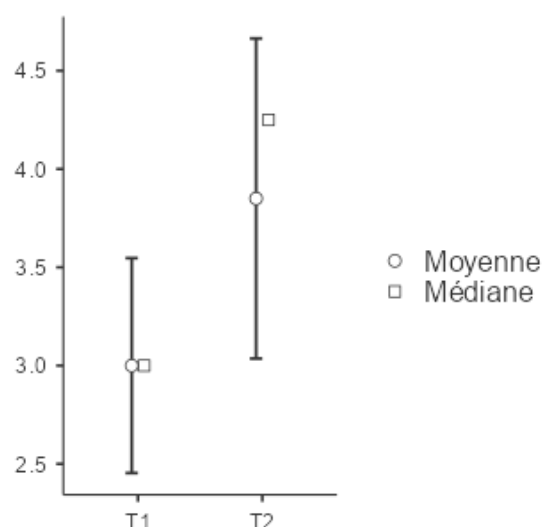


Figure 52 – Forte augmentation des scores individuels médians et moyens de B à la sous-échelle SEP(srl) sur T1-T2

2.1.3 Augmentation significative du SEP(srl) de SE sur T1-T2 et T3-T4

2.1.4 Discussion (HO 2a)

Ces résultats convergents (Tableau 10) nous permettent de valider l’hypothèse HO 2a selon laquelle le dispositif expérimental de coévaluation numérique contribue à l’augmentation du sentiment d’efficacité personnelle pour l’autorégulation des apprentissages. Ces résultats corroborent notamment les conclusions de (Hawe & Dixon, 2017; Meusen-Beekman *et al.*, 2016; Nicol & Macfarlane-Dick, 2006) relativement aux effets positifs de la coévaluation sur l’autorégulation des apprentissages.

Hypothèse	Echantillon	Intervalle	Indicateur	Résultats
HO 2a : Amélioration de l’autorégulation des apprentissages	B	T1-T2	SEP(srl)	Forte augmentation
	SE	T1-T2	SEP(srl)	Augmentation modérée
	SE	T3-T4	SEP(srl)	Faible augmentation

Tableau 10 – Synthèse des résultats obtenus pour l’hypothèse HO 2a : amélioration de l’autorégulation des apprentissages

A cet égard, l’augmentation particulièrement forte du SEP(srl), constatée pour l’échantillon B (classe de 3^e) sur l’intervalle T1-T2 constitue un fait notable dans la mesure où il s’agit d’une classe en établissement « Réseau d’éducation prioritaire » (REP). En effet, l’autorégulation des apprentissages pose des difficultés particulières pour cette classe, en lien avec des problématiques de motivation (Ryan & Deci, 2002), de volition (Cosnefroy, 2010; Fenouillet, 2009) – à fortiori pour plusieurs élèves en situation de quasi-décrochage scolaire (Cosnefroy & Fenouillet, 2009). La très forte augmentation du SEP(srl) de la classe B entre T1

et T2 semble donc confirmer les résultats de (Dignath *et al.*, 2008; Quesada *et al.*, 2019), selon lesquels la coévaluation favorise singulièrement la motivation des élèves. En s'appuyant sur les conclusions de (Meusen-Beekman *et al.*, 2016), la coévaluation numérique individualisée permettrait donc d'opérationnaliser la différenciation pédagogique en favorisant « l'engagement actif de l'élève dans la gestion des activités d'apprentissage » (Allal, 2017, p. 171) – c'est-à-dire en l'empuissantant pour qu'il prenne part à ce que (Lescouarch, 2021 : 8) nomme le « collectif apprenant ».

On constate cependant que pour l'ensemble des élèves participants la hausse observée est moindre concernant l'intervalle T3-T4. En d'autres termes, si la coévaluation numérique individualisée paraît susciter des progrès significatifs en termes d'autorégulation des apprentissages dans un premier temps, il semble néanmoins que l'amplitude de ces progrès s'amenuise par la suite. On peut supposer que cette observation illustre « l'effet plafond » décrit par (Forget, 2017b, p. 30) : les élèves acquérant une maîtrise experte eu égard aux compétences et métacompétences cibles, continuent à affiner leur degré de maîtrise, mais leur marge de progression est restreinte. Vis-à-vis des élèves qui rencontrent quant à eux des difficultés, ce processus pourrait donc ne pas susciter un accroissement notable des inégalités socioscolaires, mais participer au contraire d'une plus grande égalité des acquis. Eu égard aux compétences cibles en français langue première, ce résultat peut être mis en perspective avec la présence, établie par ailleurs, d'une dynamique globale de progrès pour toute la classe.

2.2 Métacognition (HO 2b)

Pour rappel, nous testons l'hypothèse HO 2b – c'est-à-dire « La coévaluation numérique améliore la métacognition » – à partir de plusieurs indicateurs, *e.g.* le score de coévaluateur, le score d'auto-évaluateur et le score de scripteur. Toutes les données afférentes sont numériques, exprimées en valeurs brutes. Elles sont analysées quantitativement, au moyen des tests de Wilcoxon, Friedman et Spearman, selon l'approche détaillée *supra*. Au total, 22 analyses inférentielles ont été pratiquées, dont deux ont révélé des résultats significatifs.

2.2.1 Tendances haussières des scores de coévaluateurs moyens

Les tests de Wilcoxon que nous avons appliqués aux scores individuels de coévaluateurs de tous les échantillons appariés de T1 à T6 n'ont révélé aucun résultat significatif.

Cependant lorsque l'on considère l'évolution de la moyenne des scores de coévaluateurs pour l'ensemble des élèves participants (super-échantillon SE, N = 61) sur l'intervalle T2-T4, on constate une tendance à la hausse (Figure 53).

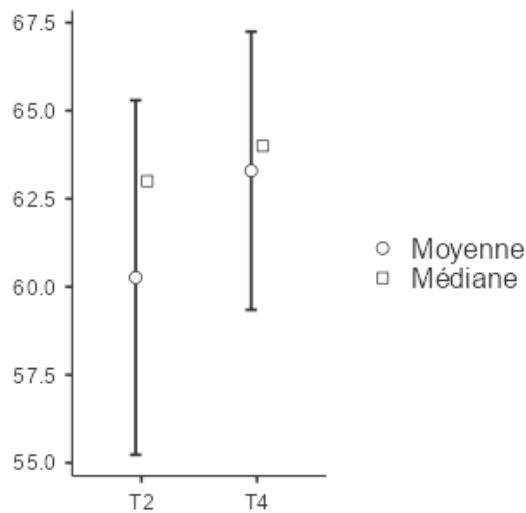


Figure 53 – Tendance haussière du score moyen de coévaluateur pour le super-échantillon SE (N = 61) sur T2-T4

De même, lorsque l'on considère l'évolution du score moyen de coévaluateur pour l'échantillon C (N = 23) sur l'intervalle T2-T6, on constate une tendance à la hausse (Figure 54).

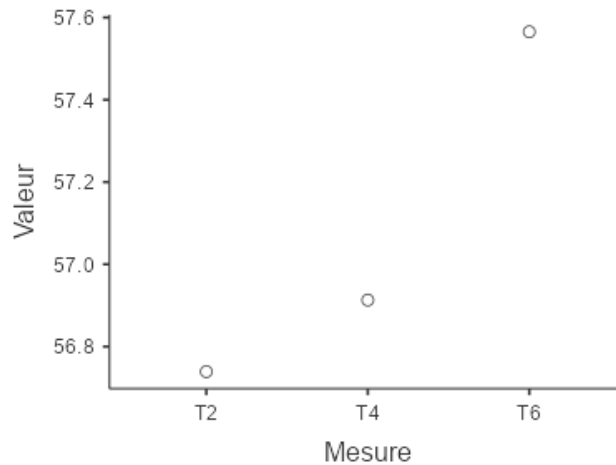


Figure 54 – Hausse du score moyen de coévaluateur pour l'échantillon C (N = 23) sur l'intervalle T2-T6

2.2.2 Corrélation entre scores de scripteurs et de coévaluateurs à T4

Les corrélations de Spearman pratiquées sur les variables « scores d'auto-évaluateurs » et « scores de scripteurs » n'ont fourni aucun résultat significatif.

Toutefois, celles pratiquées sur les variables « scores de coévaluateurs » et « scores de scripteurs » ont montré une corrélation positive modérée pour l'échantillon A à T4 ($p = 0.042$;

$r = 0.569$; $N = 13$), ainsi qu'une corrélation positive légère mais significative pour le super-échantillon SE à T4 ($p = 0.007$; $r = 0.332$; $N = 65$).

2.2.3 Discussion (HO 2b)

Nous n'avons pas pu établir une différence significative concernant les notes de coévaluateurs de l'ensemble des élèves participants entre T2 et T4 – ou entre T2 et T6 pour l'échantillon C ; cependant nous avons pu établir des tendances haussières concernant les scores moyens. Par ailleurs, les coefficients de corrélation positifs que nous avons mis en évidence signifient qu'à T4 les notes de scripteurs tendent à augmenter quand les notes de coévaluateurs augmentent – et réciproquement. Ceci est valable pour l'ensemble des élèves participants – c'est-à-dire pour le super-échantillon (SE) –, et singulièrement pour l'échantillon A.

Or il a été établi précédemment que les scores de scripteurs ont augmenté sur l'ensemble de la période expérimentale pour l'ensemble des élèves. On peut donc en déduire que, sur l'intervalle T3-T4, les scores de coévaluateurs du super-échantillon (SE) ont eux aussi augmenté. La corrélation positive concernée est faible, ce qui semble indiquer que la hausse des scores de coévaluateurs de l'ensemble des élèves est moindre que la hausse modérée des scores de scripteurs préétablie pour la période considérée. On peut toutefois affirmer la présence d'une hausse légère mais significative des scores individuels de coévaluateurs de l'ensemble des élèves entre T3 et T4 .

Bref, nous ne sommes pas en mesure de valider l'hypothèse HO 2b – c'est-à-dire « La coévaluation améliore la métacognition » – pour l'intervalle T1-T2 ; néanmoins les résultats obtenus convergent en ce qui concerne l'intervalle T3-T4 : l'hypothèse HO 2b est donc partiellement validée. Dans le Tableau 11, les lignes colorées en orange signalent les résultats qui ont pu être combinés pour aboutir à cette conclusion, représentée en vert.

Hypothèse	Echant.	T	Indicateurs	Résultats
HO 2b : Amélioration de la métacognition	SE	T2-T4	Score de coévaluateur (moyenne)	Tendance haussière
	C	T2-T6	Score de coévaluateur (moyenne)	Tendance haussière
	A	T4	Score de scripteur Score de coévaluateur	Corrélation positive modérée
	SE	T4	Score de scripteur Score de coévaluateur	Faible corrélation positive
	SE	T2-T4	Score de scripteur	Augmentation modérée
	SE	T2-T4	Score de coévaluateur	Augmentation significative

Tableau 11 – Synthèse des résultats obtenus pour l'hypothèse HO 2b : amélioration de la métacognition

Ces résultats ne corroborent donc que partiellement ceux de (Deeley, 2014; Sivan, 2000) relativement aux effets positifs de la coévaluation numérique sur les capacités métacognitives des apprenants. De ce point de vue, la contribution de la coévaluation numérique à la mise en œuvre de la différenciation pédagogique, par l'entremise des processus métacognitifs (Grangeat, 1999), reste donc à établir. Nos conclusions s'inscrivent toutefois dans le prolongement des travaux de (Kaplan, 2019) en fournissant, à l'instar de (Bouziane & Ziad, 2018a), une illustration concrète des effets d'un artéfact numérique de coévaluation en particulier sur la métacognition, ici Moodle Atelier. A cet égard, un conflit instrumental (Rabardel, 1995) relatif à l'usage de cet artéfact numérique pourrait avoir causé un déséquilibre vis-à-vis des deux autres artéfacts avec lesquels il interagit, c'est-à-dire les artéfacts didactique et pédagogique (Marquet, 2005b). Un tel conflit instrumental serait à même d'expliquer l'absence de hausse significative du sentiment d'efficacité personnelle pour les apprentissages autorégulés pendant l'intervalle T1-T2. Néanmoins, l'amélioration globale des compétences métacognitives (scores de coévaluateur) pour l'ensemble des échantillons entre T2 et T4 – autrement dit, pendant la seconde moitié de la phase expérimentale – pourrait signifier que des processus de régulation instrumentale ont eu lieu (Zeller & Marquet, 2021b).

2.3 Sentiment d'efficacité personnelle (HO 2c)

Comme détaillé *supra*, nous testons l'hypothèse HO 2c – c'est-à-dire « La coévaluation numérique améliore le sentiment d'efficacité personnelle » –, au moyen de l'indicateur « sentiment d'efficacité personnelle » (SEP), en nous appuyant sur les réponses des élèves

participants aux 37 items du questionnaire SEPE⁵⁴ ($\alpha = 0.913$; $\omega = 0.915$). Pour rappel, ces données textuelles sont collectées via une échelle de Likert à cinq points, puis transposées sous la forme de données ordinales. Nous les analysons ensuite quantitativement en tant que données à intervalle, au moyen des tests de Wilcoxon et de Friedman, pratiqués sur scores individuels médians et scores individuels moyens.

En appliquant systématiquement l'approche statistique décrite plus haut, nous avons réalisé 34 analyses inférentielles, dont 5 ont généré des résultats significatifs. Les analyses longitudinales sur échantillons appariés, c'est-à-dire les tests de Friedman sur scores individuels médians et scores individuels moyens, n'ont généré aucun résultat significatif ($p > .05$). Les analyses pair-à-pair réalisées sur échantillons appariés au moyen du test des rangs signés de Wilcoxon, sur scores individuels médians et scores individuels moyens, ont quant à elles généré des résultats significatifs, détaillés ci-après.

2.3.1 Augmentation modérée du SEP de SE sur T1-T2 et T3-T4

Pour l'ensemble des élèves participants (super-échantillon SE, $N = 46$), les tests de Wilcoxon pratiqués sur scores individuels moyens ont révélé une augmentation modérée du SEP entre T1 et T2 ($p = 0.021$; $r = 0.413$) et entre T3 et T4 ($p < 0.001$; $r = 0.611$), comme illustré dans la Figure 55 et la Figure 56.

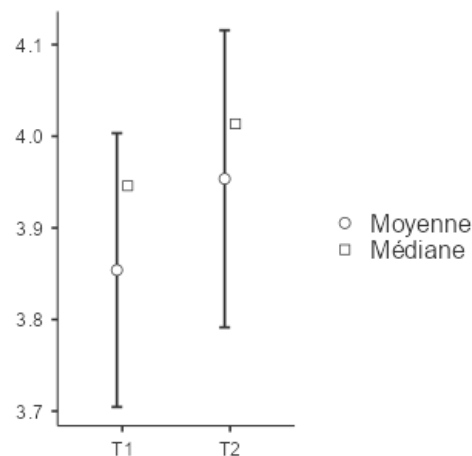


Figure 55 – Augmentation modérée des scores individuels moyens des élèves de SE à l'échelle SEP sur T1-T2

⁵⁴ SEPE : Sentiment d'efficacité personnelle des enfants, pré-adolescents et adolescents.

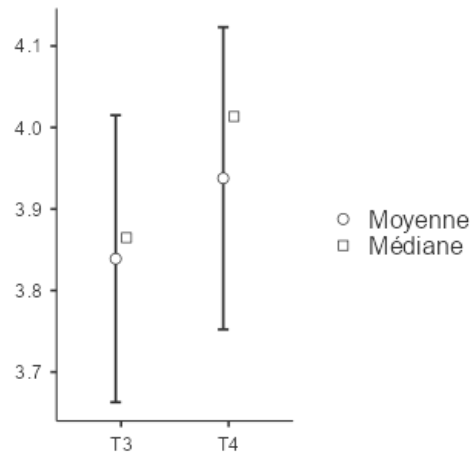


Figure 56 – Augmentation modérée des scores individuels moyens des élèves de SE à l'échelle SEP sur T3-T4

2.3.2 Augmentation modérée du SEP de A' et C sur T3-T4

Ces résultats observés concernant le super-échantillon SE ont pu être corroborés avec ceux des échantillons A' ($p = 0.049$; $r = 0.654$; $N = 12$) et C ($p = 0.041$; $r = 0.588$; $N = 16$). Pour ces deux échantillons en effet, le test de Wilcoxon sur scores individuels moyens a permis d'établir une augmentation modérée du SEP entre T3 et T4 (Figure 57 et Figure 58).

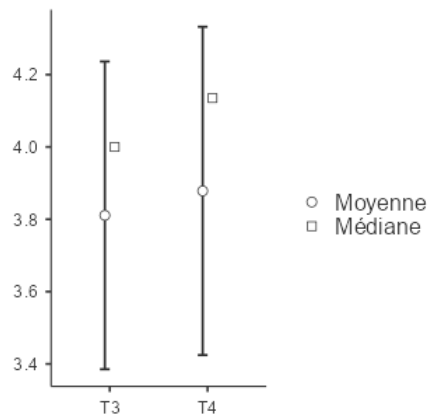


Figure 57 – Augmentation modérée des scores individuels moyens des élèves de A' à l'échelle SEP sur T3-T4

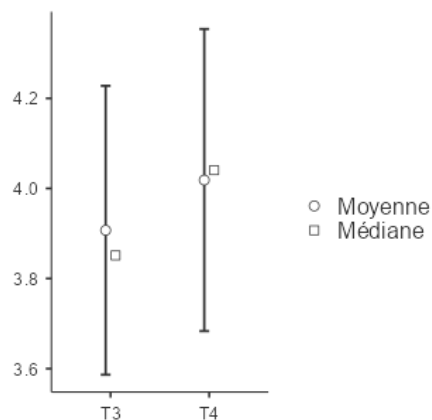


Figure 58 – Augmentation modérée des scores individuels moyens des élèves de C à l'échelle SEP sur T3-T4

2.3.3 Forte augmentation du SEP de B sur T1-T2

De surcroît, pour les élèves de l'échantillon B (N = 10), le test de Wilcoxon sur scores individuels moyens ($p = 0.01$; $r = 0.891$) a révélé une forte augmentation entre T1 et T2 (Figure 59).

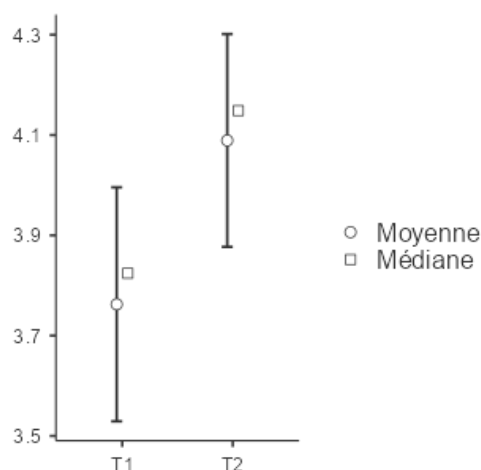


Figure 59 – Forte augmentation des scores individuels moyens des élèves de B à l'échelle SEP sur T1-T2

2.3.4 Discussion (HO 2c)

Les résultats significatifs récapitulés dans le Tableau 12 permettent de valider l'hypothèse HO 2c, c'est-à-dire « La coévaluation numérique améliore le sentiment d'efficacité personnelle ». En particulier, une hausse modérée du SEP est établie pour l'ensemble des participants – c'est-à-dire le super-échantillon (SE) – et pour toute la durée du protocole expérimental commune à tous les échantillons, c'est-à-dire de T1 à T4. Précisons que par « augmentation modérée », nous entendons une augmentation d'amplitude moyenne, établie au moyen d'un estimateur de taille d'effet – en l'occurrence, le coefficient de corrélation entre rangs bisériés – ici proche de 0.5, sachant que la valeur maximale pour une corrélation positive est 1.

Hypothèse	Echantillon	Intervalle	Indicateur	Résultat
HO 2c : Sentiment d'efficacité personnelle	B	T1-T2	SEPE	Forte augmentation
	SE	T1-T2	SEPE	Augmentation modérée
	A'	T3-T4	SEPE	Augmentation modérée
	C	T3-T4	SEPE	Augmentation modérée
	SE	T3-T4	SEPE	Augmentation modérée

Tableau 12 – Synthèse des résultats obtenus pour l'hypothèse HO 2c : amélioration du sentiment d'efficacité personnelle

Ce résultat corrobore ceux de (Boud & Falchikov, 2006; Orsmond *et al.*, 2013) concernant les effets positifs de la coévaluation sur le sentiment d'efficacité personnelle. On peut supposer à cet égard que la diminution des biais sociocognitifs au moyen d'une fonctionnalité typique de l'artéfact numérique, *e.g.* l'anonymisation en double aveugle du processus coévaluatif, contribue à désinhiber les élèves et à renforcer leur SEP (Dochy *et al.*, 1999; Topping, 1998). Surtout, compte tenu des résultats de la dernière enquête PISA (OCDE, 2019b) relativement au SEP des élèves français – significativement plus faible que la moyenne des autres pays de l'OCDE –, la coévaluation numérique semble pouvoir contribuer à améliorer la situation, tout du moins celle des collégiens.

L'absence de résultat significatif concernant l'échantillon A pose toutefois question. Là encore, on peut émettre l'hypothèse qu'un conflit instrumental (Rabardel, 1995) en lien avec l'usage de l'artéfact numérique a eu des répercussions sur l'artéfact didactique et/ou sur l'artéfact pédagogique (Marquet, 2005b), de sorte que les élèves ont moins eu le sentiment de progresser (Bandura, 2007) dans leur maîtrise des compétences cibles. A cet égard, les travaux de (Darwanto *et al.*, 2020; Lu & Law, 2012) incitent à prêter une attention toute particulière au feedback (écrits commentatifs) que ce type de dispositif est à même de générer, mais dont les élèves n'ont peut-être pas suffisamment bénéficié ici.

2.4 Discussion (HO 2)

Selon l'approche par méthodes mixtes, nous pouvons combiner les résultats que nous venons de présenter concernant l'autorégulation des apprentissages, la métacognition et le sentiment d'efficacité personnelle pour tester l'hypothèse HO 2 – c'est-à-dire « La coévaluation numérique contribue à l'autodirection des apprentissages » –, telle qu'elle est opérationnalisée dans la présente recherche.

Pour commencer, nous avons validé l'hypothèse selon laquelle la coévaluation numérique améliore l'autorégulation des apprentissages (HO 2a). Nos résultats convergent par ailleurs pour valider l'hypothèse selon laquelle la coévaluation améliore la métacognition (HO 2b). Enfin, nos résultats nous ont permis de vérifier que la coévaluation numérique améliore le sentiment d'efficacité personnelle des élèves participants (HO 2c).

D'après (Carré, 2010), la combinaison des trois conclusions précédentes nous permet donc de valider l'hypothèse selon laquelle le dispositif expérimental de coévaluation numérique individualisée contribue à favoriser l'autodirection des apprentissages pour les

élèves participants (HO 2). Le Tableau 13 récapitule le processus de combinaison des résultats qui nous a permis d'aboutir à cette conclusion.

	Hypothèses opérationnelles (HO)		Validation ?
HO 2	Amélioration de l'autodirection des apprentissages		Oui
	HO 2a	Amélioration de l'autorégulation des apprentissages	Oui
	HO 2b	Amélioration de la métacognition	Oui
	HO 2c	Augmentation du sentiment d'efficacité personnelle	Oui

Tableau 13 – Synthèse des résultats obtenus pour l'hypothèse HO 2 : Amélioration de l'autodirection des apprentissages

Les implications de cette conclusion sont significatives, eu égard à la capacité d'un dispositif tel que celui expérimenté dans la présente recherche à opérationnaliser la différenciation pédagogique. Au plan praxéologique d'une part, la contribution de la coévaluation numérique à l'autodirection des apprentissages suppose que des déterminants de la DP tels que l'autorégulation des apprentissages, la motivation, la métacognition ou le sentiment d'efficacité personnelle sont mobilisés (Allal, 2017, 2020). Au plan épistémologique d'autre part, les résultats obtenus ici confirment la pertinence des préconisations de (Forget, 2017b) relativement à l'utilisation des méthodes mixtes pour opérationnaliser un objet de recherche multivarié tel que la différenciation pédagogique.

3 Apprentissages coopératifs (HO 3)

Pour rappel, nous testons l'hypothèse HO 3 – c'est-à-dire « La coévaluation numérique favorise les apprentissages coopératifs » au moyen de deux variables opérationnelles :

- Le score individuel (médian et moyen) à l'échelle SEP(coop)⁵⁵ ;
- Le score individuel moyen à l'échelle SEC-IAP⁵⁶ (valeur brute).

Nous utilisons pour ce faire les tests suivants :

- Test de Wilcoxon (pour comparer deux temps de mesure pair-à-pair) ;
- Test de Friedman (pour comparer deux temps de mesure ou plus) ;
- Test de Spearman (pour rechercher des corrélations entre les deux indicateurs précités).

⁵⁵ SEP(coop) : Sentiment d'efficacité personnelle pour les apprentissages coopératifs.

⁵⁶ SEC-IAP : Sentiment d'efficacité collective (indice agrégatif personnel).

En appliquant systématiquement l'approche statistique détaillée *supra*, nous avons pratiqué 45 analyses inférentielles, dont dix ont généré des résultats significatifs.

3.1 SEP(coop)

3.1.1 Forte augmentation du SEP(coop) de A' et de B sur T1-T2

Pour commencer, le test des rangs signés de Wilcoxon montre une forte augmentation du SEP(coop) individuel moyen de A' ($p = 0.02$; $r = 0.944$; $N = 8$) et de B ($p = 0.019$; $r = 0.855$; $N = 10$) sur l'intervalle T1-T2. La Figure 60 et la Figure 61 illustrent ces augmentations.

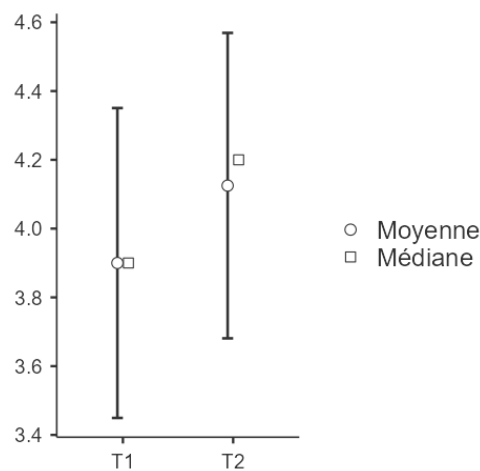


Figure 60 – Forte augmentation des scores SEP(coop) individuels moyens de A' sur T1-T2 ($N = 8$)

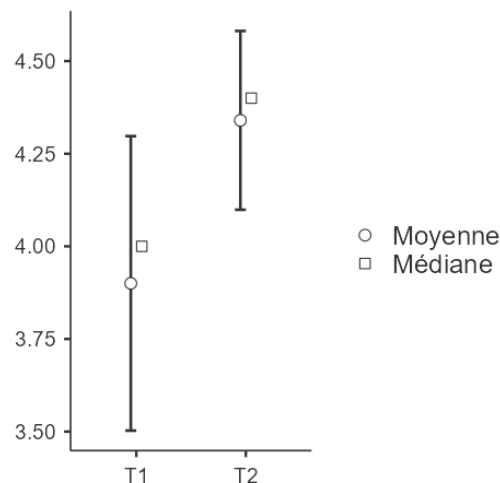


Figure 61 – Forte augmentation des scores SEP(coop) individuels moyens de B sur T1-T2 ($N = 10$)

3.1.2 Différence significative concernant le SEP(coop) de SE sur T1-T4

Par ailleurs, le test de Friedman révèle une différence très significative ($p = 0.001$) concernant les scores SEP(coop) moyens du super-échantillon ($N = 30$) sur l'intervalle T1-T4.

L'analyse post-hoc réalisée au moyen du test de Durbin-Conover confirme la différence significative constatée au moyen du test de Friedman pour les intervalles T1-T2, T1-T3 et T1-T4. Elle montre toutefois une absence de différence significative entre les scores individuels moyens de T2 et de T3.

3.1.3 Augmentation modérée du SEP(coop) de SE sur T1-T2

Enfin, le test de Wilcoxon pratiqué sur scores individuels moyens indique une augmentation modérée du SEP(coop) du super-échantillon SE (N = 46) sur l'intervalle T1-T2 ($p = 0.011$; $r = 0.478$), illustrée par la Figure 62.

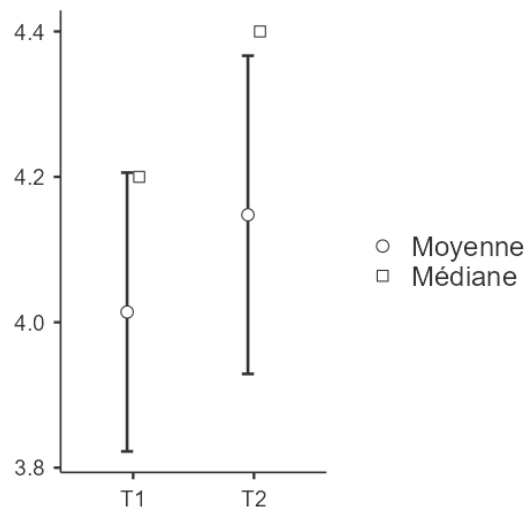


Figure 62 – Augmentation modérée du score SEP(coop) individuel moyen de SE sur T1-T2 (N = 46)

3.2 SEC-IAP

3.2.1 Très forte augmentation du SEC-IAP de A sur T3-T4

En ce qui concerne le second indicateur – c'est-à-dire le score, exprimé en valeur brute, obtenu par les élèves participants à l'item unique mesurant le sentiment d'efficacité collective selon l'indice agrégatif personnel (SEC-IAP) –, on constate une très forte augmentation pour l'échantillon A (N = 7) sur l'intervalle T3-T4 ($p = 0.048$; $r = 1$). La Figure 63 illustre ce résultat.

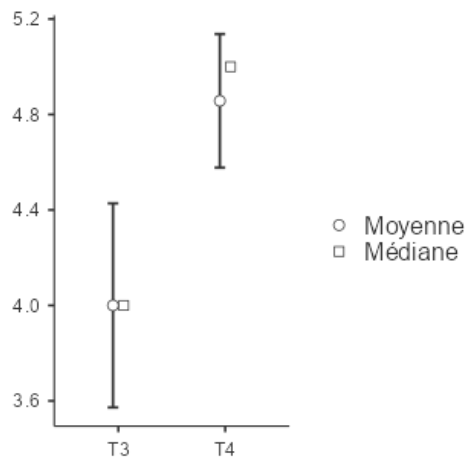


Figure 63 – Très forte augmentation du SEC-IAP de A sur T3-T4 (N = 7)

3.2.2 Forte augmentation du SEC-IAP de SE sur T3-T4

La hausse précitée est corroborée par une forte augmentation ($p = 0.008$; $r = 0.735$) du SEC-IAP sur le même intervalle pour l'ensemble des élèves participants, c'est-à-dire pour le super-échantillon SE (N = 30). La Figure 64 illustre cette augmentation.

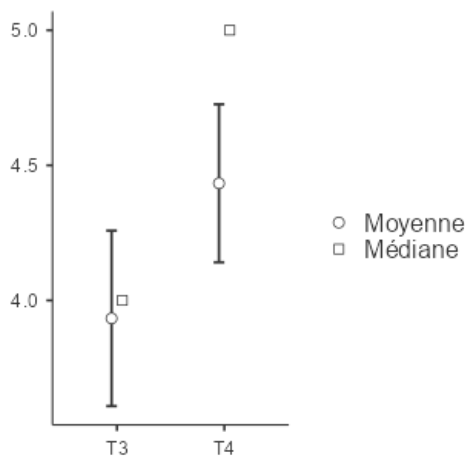


Figure 64 – Forte augmentation du SEC-IAP de SE sur T3-T4 (N = 30)

3.3 Corrélations entre SEP(coop) et SEC-IAP

Au moyen du test de Spearman, nous cherchons pour finir à consolider les résultats précités en vérifiant si ces deux indicateurs mesurent effectivement le même phénomène, c'est-à-dire s'ils entretiennent une relation de corrélation.

3.3.1 Forte corrélation positive entre SEP(coop) et SEC-IAP pour A et C à T4

On constate alors une forte corrélation positive entre les scores individuels médians à la sous-échelle SEP(coop) et les scores individuels à l'échelle SEC-IAP pour les élèves des échantillons A ($p = 0.046$; $r = 0.764$; $N = 7$) et C ($p = 0.05$; $r = 0.726$; $N = 13$) au temps de mesure T4.

3.3.2 Corrélation positive modérée entre SEP(coop) et SEC-IAP pour SE à T4 et C à T6

De la même manière, on identifie une corrélation positive modérée entre les scores individuels médians à la sous-échelle SEP(coop) et les scores individuels à l'échelle SEC-IAP du super-échantillon SE au temps de mesure T4 ($p = 0.05$; $r = 0.494$; $N = 30$) et de l'échantillon C au temps de mesure T6 ($p = 0.025$; $r = 0.466$; $N = 23$).

3.4 Discussion (HO 3)

Les dix résultats significatifs détaillés ci-avant sont résumés dans le Tableau 14.

Hypothèse	Echant.	Intervalles ou T de mesure	Indicateurs	Résultats
HO 3 : Amélioration des apprentissages coopératifs	A	T3-T4	SEC-IAP	Très forte augmentation
	SE	T3-T4	SEC-IAP	Forte augmentation
	A'	T1-T2	SEP(coop)	Forte augmentation
	B	T1-T2	SEP(coop)	Forte augmentation
	SE	T1-T2	SEP(coop)	Augmentation modérée
	SE	T1-T4	SEP(coop)	Différence significative
	A	T4	SEP(coop) médian SEC-IAP	Forte corrélation positive
	C	T4	SEP(coop) médian SEC-IAP	Forte corrélation positive
	SE	T4	SEP(coop) médian SEC-IAP	Corrélation positive modérée
	C	T6	SEP(coop) médian SEC-IAP	Corrélation positive modérée

Tableau 14 – Synthèse des résultats obtenus pour l'hypothèse HO 3 : amélioration des apprentissages coopératifs

Eu égard au SEC-IAP, nous avons pu établir une forte augmentation pour le super-échantillon (SE) sur l'intervalle T3-T4, mais nous ne disposons pas de résultats significatifs concernant l'intervalle T1-T2. En appui sur ce seul indicateur, l'hypothèse HO 3 ne serait donc que partiellement validée.

En revanche, relativement à l'autre indicateur utilisé pour tester HO 3 – c'est-à-dire le sentiment d'efficacité personnelle pour les apprentissages coopératifs –, plusieurs des résultats obtenus peuvent être combinés pour établir une augmentation significative concernant le super-échantillon (SE), pour l'ensemble de la période expérimentale, commune à tous les échantillons (autrement dit l'intervalle T1-T4). En effet, pour le super-échantillon (SE), nous avons mis en évidence une augmentation du SEP(coop) entre T1 et T2, ainsi qu'une augmentation du SEC-IAP entre T3 et T4. D'autre part, nous avons établi, pour ce même super-échantillon (SE), une corrélation positive entre le SEC-IAP et le SEP(coop) individuel médian à T4. En d'autres termes, lorsque le SEC-IAP augmente – ce qui est le cas sur l'intervalle T3-T4 –, le SEP(coop) augmente lui aussi.

De ce qui précède, on peut déduire que le SEP(coop) de SE tend à augmenter entre T1 et T4. On est de surcroît éclairé quant à la nature de la différence significative, observée par ailleurs, concernant le SEP(coop) individuel moyen des élèves du super-échantillon (SE) sur l'intervalle T1-T4 : cette différence significative s'avère *in fine* une hausse significative. Rappelons pour finir que nous avons établi une différence significative entre les scores SEP(coop) individuels moyens pour les intervalles T1-T2, T1-T3 et T1-T4, mais que nous avons constaté une absence de différence significative pour l'intervalle T2-T3. Or la coévaluation numérique n'a pas été mise en œuvre dans cet intervalle – ni dans le cadre du protocole, ni hors protocole. En creux, ces résultats corroborent ceux évoqués plus haut : pour l'ensemble des participants – autrement dit le super-échantillon (SE) –, ainsi que pour l'ensemble de la période expérimentale commune à tous les échantillons – c'est-à-dire pour les intervalles T1-T2, T3-T4 et T1-T4 –, lorsque les élèves participants pratiquent la coévaluation numérique, le SEP(coop) tend à augmenter, tandis qu'il ne varie pas significativement lorsque les élèves ne pratiquent pas la coévaluation numérique.

La combinaison de ces résultats nous permet de valider l'hypothèse HO 3 – c'est-à-dire « La coévaluation numérique favorise les apprentissages coopératifs ». Le Tableau 15 récapitule les combinaisons de résultats (lignes orange), qui permettent d'aboutir aux déductions précitées (lignes vertes).

Hypothèse	Echant.	Intervalles ou T de mesure	Indicateurs	Résultats
HO 3 : Apprentissages coopératifs	SE	T1-T2	SEP(coop)	Augmentation modérée
	SE	T3-T4	SEC-IAP	Forte augmentation
	SE	T4	SEP(coop) médian SEC-IAP	Corrélation positive modérée
	SE	T3-T4	SEP(coop)	Augmentation significative
	SE	T1-T2 T1-T4	SEP(coop)	Différence significative
	SE	T2-T3	SEP(coop)	Absence de différence significative
	SE	T1-T4	SEP(coop)	Augmentation significative

Tableau 15 – Combinaisons de résultats établissant la hausse du SEP(coop) de SE sur T1-T4

En appui sur ces résultats concernant le sentiment d'efficacité personnelle pour les apprentissages coopératifs, la coévaluation numérique contribuerait donc, par l'entremise de la coopération entre élèves, à la mise en œuvre de la différenciation pédagogique (Allal, 2017; Forget, 2017b).

A cet égard, nos conclusions montrent que la coévaluation numérique constitue une approche viable pour opérationnaliser conjointement pratiques coopératives et évaluation (Connac & Cieutat, 2021). En l'occurrence, bien que la coévaluation numérique soit une activité qui repose, non pas sur une interdépendance forte, autrement dit mesurable au moyen de l'indice agrégatif holistique, mais sur une interdépendance faible, c'est-à-dire mesurable par l'indice agrégatif personnel, les hausses importantes du SEC-IAP, constatées pour le super-échantillon (SE) dans l'intervalle T3-T4 – donc pour la seconde moitié de la période expérimentale commune à tous les échantillons –, indiquent que la coévaluation numérique individualisée est une activité authentiquement coopérative, dans laquelle les apprenants sont pleinement associés au processus évaluatif (Dochy *et al.*, 1999).

Nous avons pourtant vu *supra* que le dispositif de coévaluation numérique expérimenté ici est hybride : bien que la situation d'enseignement/apprentissage soit présentielle – c'est-à-dire en salle informatique –, elle est mise en œuvre par ordinateurs interposés avec, de surcroît, l'asynchronie qu'impliquent les différences inter-individuelles en termes de rythme de travail. Ce dispositif semble néanmoins contribuer de manière partielle mais significative à une distribution des enjeux du processus évaluatif entre l'enseignant et les apprenants (Blau & Shamir-Inbal, 2018; Quesada *et al.*, 2019; Reynaud, 2022). La coévaluation numérique

individualisée, telle qu'elle a été mise en œuvre dans ce protocole de recherche, contribuerait ainsi à la mise en œuvre d'un collectif apprenant au sens de (Lescouarch, 2021).

Singulièrement, la très forte augmentation du sentiment d'efficacité collective mesurée par l'indice agrégatif personnel (SEC-IAP) constatée entre T3 et T4 pour l'échantillon A, ainsi que – plus globalement – la forte augmentation constatée pour le super-échantillon (SE) sur le même intervalle, permettent de proposer deux autres pistes interprétatives.

Pour commencer, l'absence de résultat significatif en termes de SEC-IAP pour l'intervalle T1-T2 laisse supposer que l'expérience en matière de coévaluation numérique est en l'occurrence un déterminant important de l'indice agrégatif personnel du sentiment d'efficacité collective. Cette interprétation accrédirait l'hypothèse selon laquelle les apprenants bénéficient d'autant plus de la coévaluation numérique qu'ils la pratiquent souvent. D'une part, cette hypothèse fait écho à un des résultats précédemment établis, à savoir : le non-usage de la coévaluation numérique s'accompagne d'une absence d'augmentation du sentiment d'efficacité personnelle pour les apprentissages coopératifs. D'autre part, cela corroborerait les résultats de (Quesada *et al.*, 2019) concernant le problème que pose le manque de continuité avec ce type de pratiques évaluatives. A cet égard, si l'usage d'un artéfact numérique pour pratiquer la coévaluation est susceptible de générer des conflits instrumentaux (Rabardel, 1995), causant une perturbation des apprentissages (Marquet, 2005b), la résolution concomitante de ces conflits, au moyen de processus de régulation instrumentale mis en œuvre par les usagers (Zeller & Marquet, 2021b), s'avèrerait d'autant plus importante que la continuité dans l'usage en dépend – à fortiori dans la mesure où la coévaluation individualisée n'est pas envisageable sans le recours à un artéfact numérique (Braun, 2003; Dochy *et al.*, 1999; Mostert & Snowball, 2013; Segers *et al.*, 2003a; K. Topping, 1998a).

4 Conflits et régulations instrumentaux (HO 4)

Pour rappel, nous testons l'hypothèse 4 – c'est-à-dire « La coévaluation numérique génère des conflits instrumentaux et permet de les réguler » –, au moyen de plusieurs types de données textuelles, auprès des élèves participants d'une part, et auprès des enseignants partenaires d'autre part. Lorsqu'elles sont issues de questions fermées et collectées via une échelle de Likert, ces données textuelles sont transposées sous forme de données ordinales puis analysées quantitativement en tant que données à intervalle. Lorsque les données proviennent de questions ouvertes, d'entretiens semidirigés ou de focus groupes, elles sont traitées

qualitativement : par analyse catégorielle de contenus, par analyse textométrique ou à la manière d'un récit de vie. Le détail de ces différentes modalités de traitement est fourni *supra*. Dans le présent chapitre, pour tenir compte de la diversité des données en termes de publics cibles et de modalités d'enquête, nous présentons dans un premier temps les résultats afférents au point de vue des élèves, puis ceux relatifs au point de vue des enseignantes partenaires. Après avoir abordé successivement les résultats issus des différentes modalités d'enquête utilisées, chacune de ces sections est suivie d'une discussion. Le détail des analyses (*e.g.* statistiques descriptives, graphes, verbatim d'élèves) se trouve en Annexes 1.5.

4.1 Situations d'apprentissage (HO 4a)

4.1.1 Questionnaire en ligne

Comme indiqué *supra*, trois modalités d'enquête permettent d'outiller le recueil des données textuelles auprès des élèves participants, afin de tester l'hypothèse HO 4 – c'est-à-dire « La coévaluation numérique génère des conflits instrumentaux et permet de les réguler » :

- un questionnaire en ligne composé :
 - d'un item fermé (échelle de Likert à cinq points) ;
 - de deux items ouverts ;
- un retour écrit sur expérience semidirigé ;
- deux focus groupes.

Aux temps de mesure T2 (N = 59) et T4 (N = 57), à la fin du questionnaire SEPE en ligne, les trois items destinés à tester HO 4 sont les suivants :

- « La coévaluation avec Moodle Atelier t'a-t-elle paru difficile ? » (item 1) ;
- « Pourquoi ? » (item 2) ;
- « Pour finir, as-tu des remarques ou des commentaires à faire ? » (item 3).

Relativement à l'item 1, aucune des analyses inférentielles détaillées *supra* n'a fourni de résultat significatif ($p > .05$). Les statistiques descriptives issues de l'item 1 révèlent cependant, à T2 comme à T4, des réponses situées légèrement au-dessus de 4 sur 5 en moyenne, ce qui correspondait au descriptif « La coévaluation avec Moodle Atelier m'a paru facile ». L'item 2 (à T2 et T4) et l'item 3 (à T4) font l'objet d'une analyse thématique préliminaire. Comme précisé *supra*, l'item 3 n'est proposé qu'une seule fois, au temps de mesure T2. Sur 57 élèves interrogés, 27 y ont répondu. Les résultats de ces analyses préliminaires sont consultables en Annexes 1.5 (Figure 97, Figure 98 et Figure 99).

A partir des thématiques généralistes qui émergent de ces analyses préliminaires, on procède à une seconde analyse catégorielle de contenus, artéfact par artéfact, selon l'approche instrumentale en environnement informatisé pour l'apprentissage humain (EIAH). Nous recatégorisons ainsi les différentes thématiques et verbatim d'élèves afférents, selon qu'ils concernent respectivement les artéfacts didactique, pédagogique ou technique. Comme indiqué *supra*, les verbatim d'élèves ont été toilettés/redressés. Des verbatim illustrant chaque thématique sont présentés en Annexes 1.5.4.

4.1.1.1 Artéfact didactique

Relativement à l'artéfact didactique, la principale thématique qui émerge de l'analyse catégorielle de contenus est la notion de prérequis didactique. Les élèves qui déclarent posséder une bonne maîtrise des compétences cibles déclarent également que la coévaluation a été plutôt facile pour eux (Tableau 48). A contrario, d'autres élèves déclarent que sans ce prérequis didactique il leur est difficile de s'engager dans l'activité de coévaluation (Tableau 49). En l'occurrence, environ un cinquième des élèves interrogés évoquent la difficulté à évaluer la production écrite d'un pair. Singulièrement, ils disent être confrontés à un dilemme lorsqu'il s'agit de déterminer quel est le juste niveau de maîtrise, relativement à telle ou telle compétence à coévaluer dans la production écrite d'un pair (Tableau 50).

L'analyse catégorielle de contenus permet également de mettre en évidence le problème posé à certains coévaluateurs par la qualité linguistique insuffisante de certaines productions écrites. Selon eux en effet, certains textes de pairs sont difficiles à coévaluer en raison d'un trop grand nombre d'erreurs linguistiques, qui les rendent relativement inintelligibles ou qui rendent la tâche coévaluative trop exigeante (Tableau 51).

4.1.1.2 Artéfact pédagogique

Relativement à l'artéfact pédagogique, la première thématique identifiée concerne l'appétence des élèves interrogés pour la coévaluation en soi, c'est-à-dire pour l'artéfact pédagogique lui-même (Tableau 52). Eu égard à l'acte coévaluatif, certains élèves évoquent toutefois leur peur de commettre une injustice (Tableau 53). Certains élèves déclarent par ailleurs qu'ils ont peur de ne pas évaluer comme l'enseignant. Ils évoquent à cet égard une divergence de point de vue entre leur inter-évaluation – autrement dit leur évaluation d'un pair – et l'hétéro-évaluation de l'enseignant. En outre, les verbatim d'élèves afférents suggèrent que ce conflit instrumental repose sur une vision plutôt subjective du geste évaluatif, considéré comme un acte arbitraire (Tableau 54).

En lien avec cette thématique, on peut également noter chez certains élèves un relatif manque d'expérience en matière d'inter- ou de coévaluation, associé à une difficulté à formuler un écrit commentatif constructif et argumenté (Tableau 55). A ce titre, plusieurs élèves mentionnent l'appui important qu'a représenté pour eux l'expérience acquise en matière de coévaluation (Tableau 56), singulièrement au cours de la phase expérimentale – c'est-à-dire entre T1 et T4 (ou entre T1 et T6 pour l'échantillon C). Pour autant, certains élèves ne semblent pas disposer du prérequis pédagogique nécessaire pour comprendre la signification ou l'utilité de la coévaluation. Ces élèves commettent un contresens pédagogique en considérant que la coévaluation – tâche complexe mobilisant la métacognition – est une tâche simple, avec laquelle on ne peut pas se tromper, *e.g.* parce qu'il s'agit d'un « QCM sans mauvaises réponses » (Tableau 57).

4.1.1.3 Artéfact technique

Relativement à l'artéfact technique, l'analyse catégorielle a permis d'isoler six sous-catégories qu'on peut regrouper sous la terminologie d'étayage pédagonumérique : la facilité d'utilisation, l'appétence pour l'artéfact technique en soi, la différenciation pédagogique résultant d'une horizontalisation, structurellement induite par le dispositif pédagonumérique, de la posture enseignante, l'anonymat et enfin la clarté des consignes.

Pour commencer, un nombre significatif d'élèves soulignent la simplicité d'utilisation de l'artéfact technique selon eux (Tableau 58). A cet égard, les élèves interrogés soulignent parfois tout simplement leur appétence pour l'informatique en soi (Tableau 59). Toutefois, les élèves déclarent également que les interfaces associées à l'usage de Moodle Atelier – *e.g.* clavier, correcticiel – peuvent faire l'objet de mésusages ou de non-usages générant des erreurs de langue. Au dire de certains élèves, ces erreurs linguistiques peuvent parfois abonder au point de gêner la lecture – et donc la coévaluation – des écrits de pairs (Tableau 60).

Les élèves signalent également l'importance pour eux de l'aide qu'ils ont pu recevoir de la part de leur enseignant. Aux dires des élèves, cette aide peut prendre deux formes. D'une part il peut s'agir d'une aide directe, présente et synchrone, dans le cadre d'une remédiation individualisée immédiatement effectuée en classe par l'enseignant. D'autre part l'aide peut intervenir de manière indirecte, distancielle et asynchrone – notamment, d'après les élèves interrogés, par le truchement de la fonctionnalité « Exemples de travaux » (Tableau 61).

Les élèves soulignent d'ailleurs le fait que le dispositif pédagonumérique est propice à une bonne compréhension des consignes, singulièrement grâce à la fonctionnalité « Exemples

de travaux » précitée. A ce propos, plusieurs élèves insistent sur le fait que cela leur a significativement facilité la tâche (Tableau 63).

Plusieurs élèves participants soulignent par ailleurs la valeur, selon eux protectrice et désinhibante, de la fonctionnalité de Moodle Atelier qui permet de préserver l'anonymat, à la fois pour l'évaluateur et pour l'évalué (anonymat en double aveugle).

4.1.2 Retours écrits sur expérience (classe C, T6)

Comme indiqué *supra*, à l'issue de la troisième itération du protocole expérimental (T6) avec l'échantillon C (classe de 5^e), un retour écrit sur expérience est recueilli sous la forme d'un courriel adressé par chaque élève (N = 23) à l'enseignante C, via la messagerie interne de l'environnement numérique de travail (ENT) du collège. Pour rappel, le contenu de ces courriels est semidirigé, avec deux thématiques proposées : aspects positifs du dispositif versus aspects négatifs.

Afin d'avoir une première idée de la nature du corpus, nous générons un nuage de mots pour le corpus considéré dans son ensemble, tous types de retours sur expérience confondus (positifs et négatifs). Nous ne conservons pour ce faire que les formes pleines (pas les articles, ni les pronoms et autres mots-outils) et nous retirons les verbes d'appréciation (*e.g.* aimer ou trouver) car ils sont sur-représentés en raison de la nature du corpus, et biaisent les résultats. Comme explicité *supra*, nous intégrons par ailleurs des tirets bas entre certains vocables pour préserver le cotexte, c'est-à-dire la syntaxe et les aspects sémantiques afférents. Le nuage de mots ainsi obtenu est le suivant (Figure 65) :

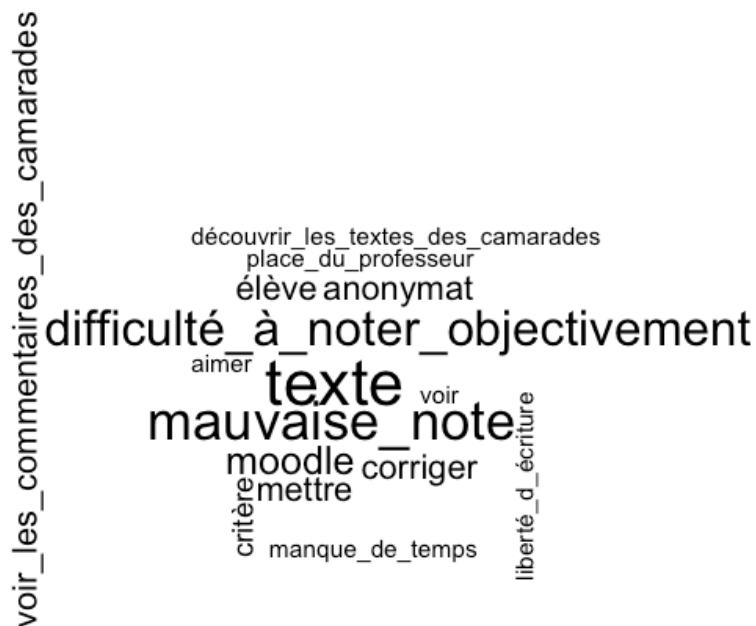


Figure 65 – Nuage de mots pour le corpus global « Retours écrits sur expérience », classe C, T6 (N = 23)

On voit ainsi apparaître plusieurs thématiques clés, *e.g.* la difficulté à noter objectivement ou la notion de « mauvaise note ». L’artéfact numérique est souvent mentionné, de même que l’anonymat (en double aveugle) qu’il permet. Les élèves mentionnent souvent le fait de voir les écrits commentatifs des camarades ; toutefois on ne peut pas savoir à ce stade si cela est pour eux mélioratif ou péjoratif. On peut faire la même remarque pour le fait de découvrir les productions écrites des camarades/pairs. Enfin, la liberté d’écriture et le manque de temps sont également mentionnés par les élèves.

Nous subdivisons ensuite le corpus global en deux sous-corpus (retours positifs vs. retours négatifs), et nous générons dans IRaMuTeQ un nuage de mots pour chacun. La Figure 66 et la Figure 67 représentent les deux nuages de mots ainsi obtenus.

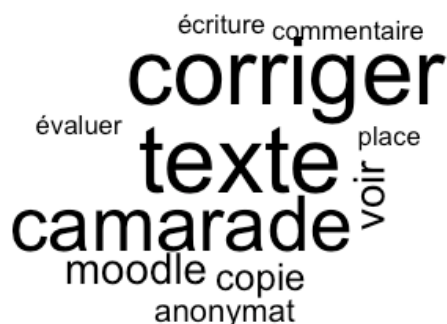


Figure 66 – Nuage de mots pour le sous-corpus « Retours positifs », classe C, T6 (N = 23)

texte mettre
 élève
note
 temps corriger
 mauvais

Figure 67 – Nuage de mots pour le sous-corpus « Retours négatifs », classe C, T6 (N = 23)

En les comparant, on constate la prévalence de deux champs lexicaux bien distincts : celui de la coévaluation pour le sous-corpus « Retours positifs », principalement avec des formes lemmatisées comme « corriger », « texte » ou « camarade », mais aussi avec des formes comme « Moodle », « anonymat » ou « commentaire », qui prennent en raison de la nature du sous-corpus une connotation méliorative. Concernant le sous-corpus « Retours négatifs », le champ lexical mis en relief est celui de la notation avec des occurrences comme « note », « mettre », « mauvais » ou « élève », qui ont ici une connotation péjorative. Toutefois, un nuage de mots ne permet pas de mettre en évidence les liens sémantiques entre les mots ou occurrences.

Nous avons pratiqué deux analyses de similitude (ADS), respectivement pour les sous-corpus « Retours positifs » et « Retours négatifs », qui nous permettent de mieux appréhender ces liens sémantiques. Les indices sur les arêtes ou branches des arbres générés nous renseignent sur le nombre de co-occurrences, c'est-à-dire lorsque les deux mots ainsi reliés se trouvent en présence l'un de l'autre dans le corpus. Grâce à un concordancier remplaçant chaque occurrence au sein de son cotexte d'appartenance, l'ADS pratiquée sur le corpus « Retours positifs » (Figure 68) permet de mettre en évidence que, selon les élèves interrogés, les aspects positifs du dispositif sont liés au fait de pouvoir « corriger » / « évaluer » les « copies » de leurs « camarades », en se mettant à la « place » du professeur et en ayant l'opportunité de « voir » le « texte » d'autrui (quatre co-occurrences). L'accent est également mis sur l'importance de l'artéfact numérique « Moodle [Atelier] » (cinq co-occurrences avec la forme « corriger »), notamment en raison de « l'anonymat » qu'il rend possible (quatre co-occurrences), et parce que la coévaluation y implique l'adjonction d'un écrit commentatif (« commentaire »). A contrario, l'ADS pratiquée sur le corpus « Retours négatifs » (Figure 69) permet de soulever la difficulté liée à la notation : soit lorsque l'élève estime avoir reçu une « mauvaise note », soit lorsqu'il s'estime lui-même injuste lorsqu'il s'agit de coévaluer la production écrite d'un pair.

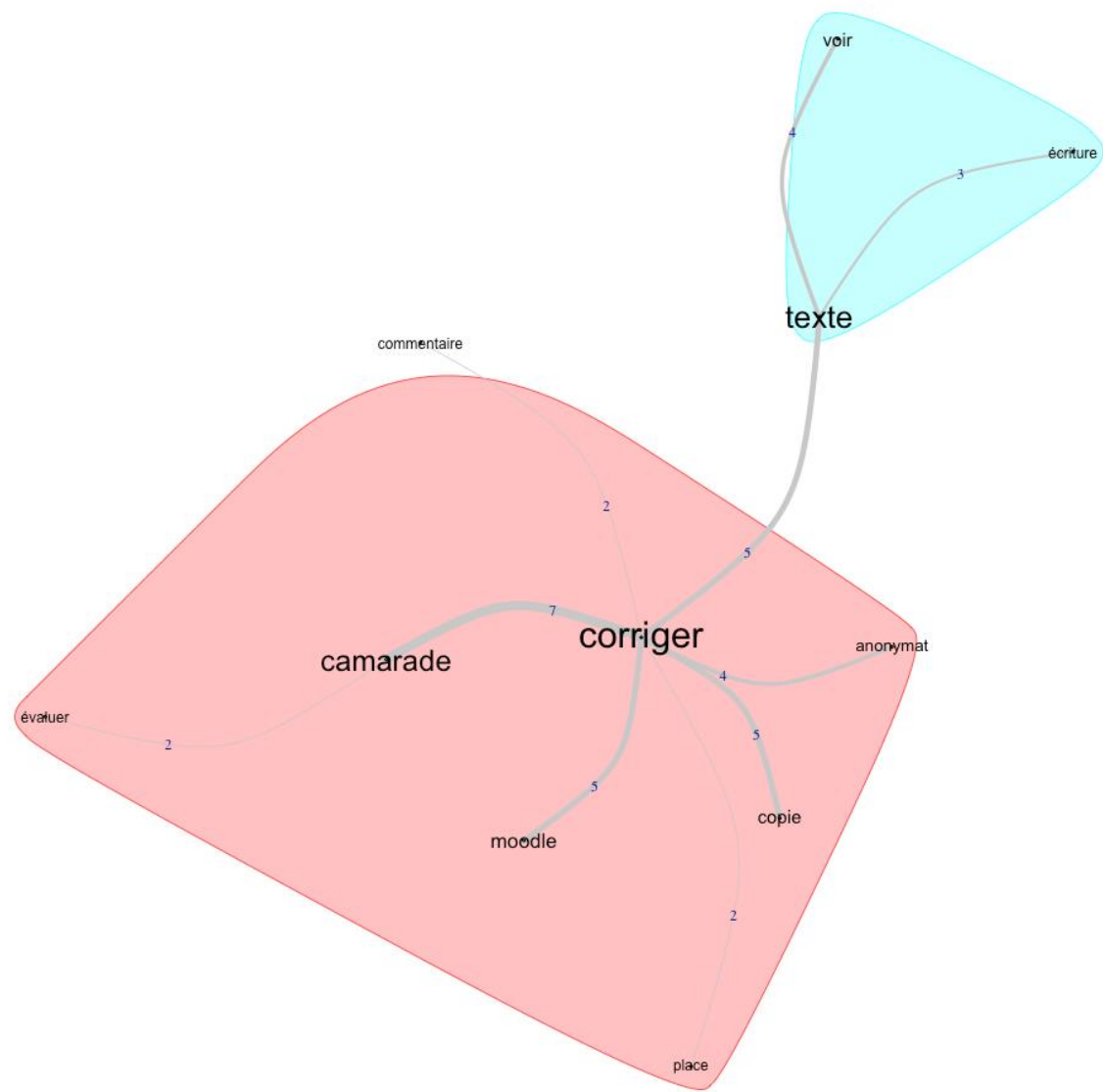


Figure 68 – Graphe de similitude pour le sous-corpus « Retours positifs »

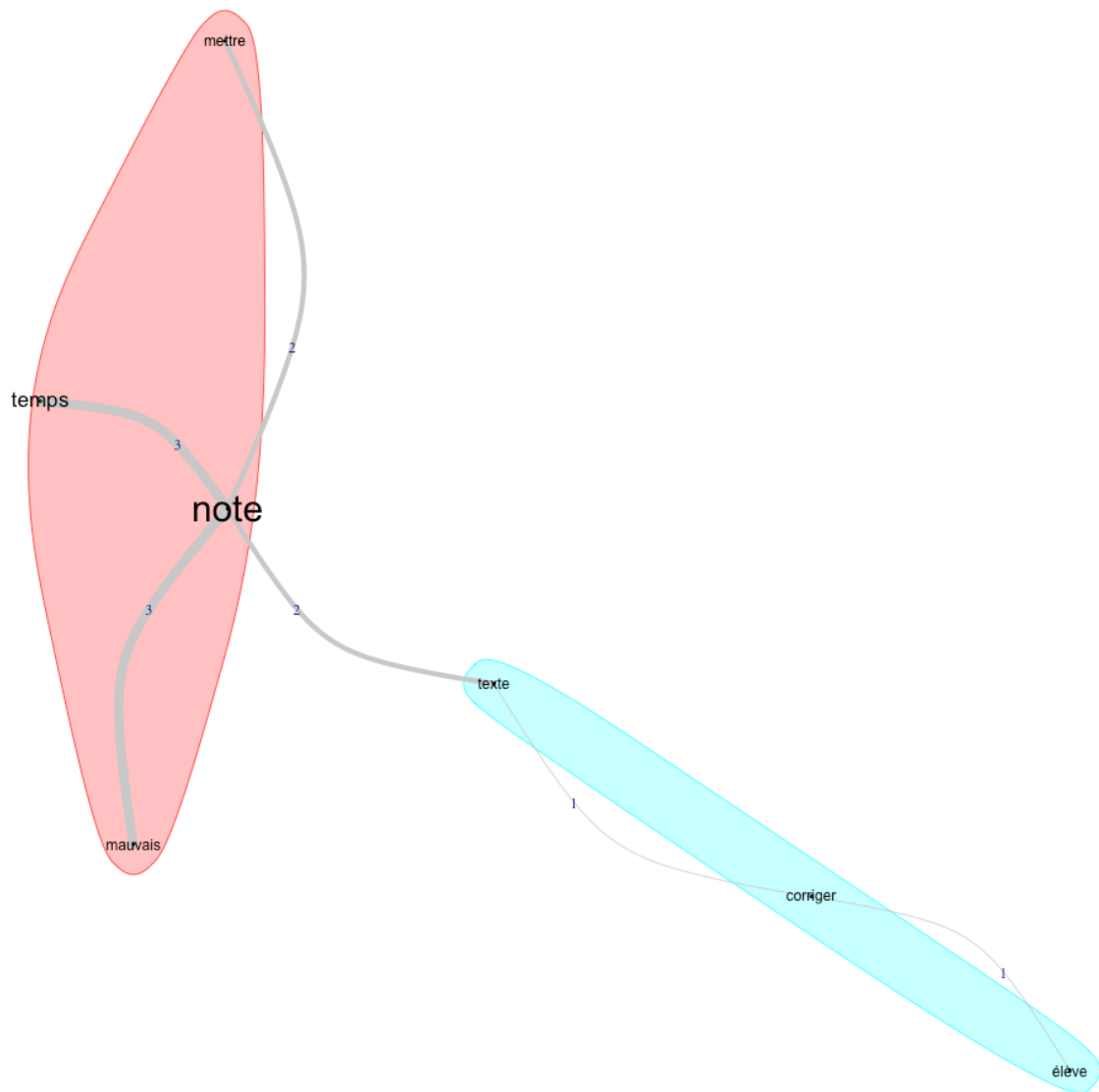


Figure 69 – Graphe de similitude pour le sous-corpus « Retours négatifs »

4.1.3 Focus groupes (classe C, T6)

Comme indiqué *supra*, deux focus groupes (N = 12 et N = 11) sont organisés avec les élèves participants de l'échantillon C (classe de 5^e), en l'absence de leur professeure, dans le prolongement de la troisième itération du protocole expérimental (T6). Pour rappel, nous pratiquons une analyse catégorielle de contenus, grâce à laquelle plusieurs thématiques sont mises en évidence.

La thématique qui apparaît ici comme la plus prégnante est celle de l'anonymat. D'après les dires des élèves dans ces focus groupes, l'anonymat est protecteur : non seulement parce qu'il permet d'éviter les conflits entre élèves coévaluateurs et élèves coévalués, mais aussi parce qu'il permet d'éviter les biais socio-affectifs – c'est-à-dire faire du favoritisme,

régler ses comptes ou faire pression. Toutefois, certains élèves déplorent que les écrits commentatifs de leurs pairs soient parfois insuffisamment explicites : pour permettre un échange entre évaluateur et évalué, ils proposent de lever l’anonymat, mais uniquement à posteriori, et en conditionnant éventuellement cela au consentement mutuel des élèves intéressés.

Ce besoin d’échanges entre évaluateur et évalué prend racine dans la thématique – également très présente dans les focus groupes – de la difficulté à évaluer la production écrite d’un pair : soit parce qu’il y a trop de critères – autrement dit trop de micro-compétences – à évaluer, soit par manque de temps alloué à la coévaluation proprement dite, soit parce que l’élève évaluateur ressent de la culpabilité vis-à-vis d’un pair dont il juge que le texte est beau, mais auquel il sera contraint d’attribuer des niveaux de maîtrise peu élevés, car il ne respecte pas suffisamment les consignes d’écriture. Ce dilemme conduit les élèves à proposer la mise en place d’un moyen de valoriser l’effort fourni par les auteurs de ces productions écrites singulières, via l’attribution de « points bonus » et d’un écrit commentatif ad hoc, non conditionnés par les critères de réussite originels.

Une troisième thématique souvent mise en avant a trait à la mise en place d’une boucle formative de rétroaction. Pour les élèves, il s’agirait ici de permettre des échanges constructifs et argumentés entre les élèves coévaluateurs et les élèves coévalués qui le souhaitent, dans le but d’affiner la coévaluation – *e.g.* pour que l’élève coévalué puisse signaler à l’élève coévaluateur, preuve à l’appui, un élément de sa production écrite qui lui aurait échappé, et qui serait susceptible d’améliorer sa note. A la faveur de débats très nourris (Annexes 5), les élèves reconnaissent toutefois la nécessité d’anonymiser ces échanges par l’entremise de l’artéfact numérique.

En lien avec ce qui précède, certains élèves manifestent leur mécontentement vis-à-vis de certaines notes de scripteurs reçues dans le cadre du dispositif, et qu’ils jugent décevantes. Les élèves concernés estiment que leur production écrite, injustement évaluée selon eux, aurait pu bénéficier d’un « droit de réponse ». Cette thématique concerne également les notes données : en effet, certains élèves se disent préoccupés par le fait que le mode de calcul de la note de coévaluateur dans Moodle Atelier est susceptible de faire baisser illégitimement leur moyenne – quand bien même ils auraient évalué avec justesse –, au cas où plusieurs autres élèves coévaluateurs se trompent lourdement et font pencher la balance de manière erronée.

Une cinquième thématique, présente dans ces focus groupes, fait écho à toutes celles précédemment évoquées : il s’agit de la notion d’écrit commentatif. En premier lieu, les élèves évoquent le caractère parfois insuffisamment explicite de cette appréciation écrite

personnalisée, qui accompagne la coévaluation dans Moodle Atelier. L'écrit commentatif est également associé par les élèves au souhait de valoriser les travaux atypiques mais méritants. Un autre type d'écrit commentatif est évoqué : celui qui pourrait être fourni dans le cadre de la boucle formative de rétroaction précédemment évoquée, entre les élèves coévaluateurs et les élèves coévalués qui y consentent. Les élèves participants insistent particulièrement sur le fait que tout écrit commentatif doit être anonyme, constructif (*e.g.* pas un commentaire laconique et stérile de type « c'est nul ») et argumenté.

Enfin, une appétence pour la coévaluation en soi est exprimée par un nombre significatif d'élèves, notamment en raison du plaisir lié, selon leurs dires, à la découverte des productions écrites des pairs, sources d'inspiration. Les élèves participants soulignent également leur goût pour le fait de donner leur avis sur la production écrite d'un pair, de dire ce qui y a plu ou déplu, et de se mettre pour ce faire « à la place du professeur ».

4.1.4 Focus groupe (classe C', T4-T5)

Selon les modalités précisées *supra*, un focus groupe est mis en œuvre entre T4 et T5 avec les élèves, non-participants à l'expérimentation, de la classe coopérative C' du collège C. Les thématiques émergent de l'analyse catégorielle pratiquée ont trait à la motivation d'une part, et à la métacognition d'autre part.

En termes de motivation, ces élèves de 6^e indiquent d'emblée qu'il leur semble « cool de faire le prof ». Dans le dispositif de coévaluation numérique qui leur est présenté, ils apprécient également à priori l'opportunité, entrevue, d'obtenir une note supplémentaire – en l'occurrence la note de coévaluateur –, car elle est d'après eux susceptible d'augmenter leur moyenne trimestrielle. Ils soulignent par ailleurs le fait que l'outil informatique, en l'occurrence le traitement de texte, permet de bénéficier d'une aide orthographique et morphosyntaxique grâce au correcticiel intégré.

Enfin, concernant la dimension métacognitive du dispositif de coévaluation numérique, les élèves de 6^e interrogés identifient d'emblée « qu'en évaluant des copies de copains de classe, on apprend à mieux repérer ses propres erreurs ».

4.1.5 Discussion (HO 4a)

Globalement, l'analyse des données recueillies auprès des élèves montre que le dispositif expérimental de coévaluation numérique génère effectivement des conflits instrumentaux (Rabardel, 1995) affectant les situations d'apprentissage, et que ces derniers se

rapportent toujours à une *combinatoire conflictuelle*, c'est-à-dire à des interactions problématiques entre plusieurs artefacts du triple enchâssement en environnement informatisé pour l'apprentissage humain : artefacts didactique, pédagogique et/ou numérique (Marquet, 2005b). On constate également que ces combinaisons conflictuelles peuvent concerner les élèves scripteurs comme les élèves coévaluateurs, et sont susceptibles d'entraîner des retards et des renoncements partiels ou totaux d'apprentissages. A contrario, ces conflits instrumentaux peuvent être accompagnés de facteurs de régulation instrumentale, pouvant eux aussi se combiner pour générer des *synergies régulatrices*, autrement dit des combinaisons de facteurs de régulation instrumentale, impliquant plusieurs types d'artefacts. Les synergies régulatrices observées dans le cadre de la présente recherche semblent à même de compenser partiellement – voire de résoudre – les conflits instrumentaux identifiés auprès des élèves participants, ce qui confirmerait la théorie de (Zeller & Marquet, 2021b).

Une typologie peut ainsi être inférée de nos résultats. Il s'agit d'une part de préciser les relations causales à l'œuvre dans le cadre des genèses instrumentales étudiées : est-on en présence d'une source de conflit instrumental ? d'un artefact affecté par un conflit instrumental ? d'un facteur de régulation ou – dans le cas de l'artefact numérique – d'une fonctionnalité régulatrice ? Dans le prolongement des travaux de (Marquet & Coulibaly, 2011; Zeller & Marquet, 2021a) notamment, on peut par ailleurs distinguer plusieurs *degrés de conflictualité*, selon que le conflit instrumental entraîne pour l'élève un retard, un renoncement partiel ou un renoncement total à la tâche d'apprentissage. Enfin, relativement aux combinaisons conflictuelles et aux synergies régulatrices, observées dans le triple enchâssement artefactuel en EIAH qui caractérise cette situation expérimentale, quels sont les artefacts concernés ?

4.1.5.1 Combinatoires artefactuelles

Pour commencer, le Tableau 16 récapitule les différents conflits instrumentaux observés, et les combinaisons artefactuelles auxquelles ils se rapportent.

Conflits instrumentaux observés	Artéfacts concernés		
Problèmes de discernement docimologique	Artéfact didactique	Artéfact pédagogique	
Inhibitions socioaffectives	Artéfact didactique	Artéfact pédagogique	
Inter- vs hétéro-évaluation : vision conflictuelle		Artéfact pédagogique	Artéfact technique
Difficulté à formuler un écrit commentatif adapté	Artéfact didactique	Artéfact pédagogique	
Contresens pédagogique à l'égard de la coévaluation	Artéfact didactique	Artéfact pédagogique	
Mésusages ou non-usage des interfaces	Artéfact didactique	Artéfact pédagogique	Artéfact technique
Absence de prérequis didactiques	Artéfact didactique	Artéfact pédagogique	
Erreurs orthographiques dans les écrits	Artéfact didactique	Artéfact pédagogique	Artéfact technique

Tableau 16 – Conflits instrumentaux observés et artéfacts afférents (violet = artéfact pédagogique ; orange = artéfact technique ; bleu = artéfact didactique)

On constate que les trois composantes artéfactuelles du dispositif peuvent être à la source de conflits instrumentaux. Plus précisément, il s'avère que l'ensemble des conflits instrumentaux observés se rapportent à plusieurs artéfacts. Le plus souvent il s'agit d'un lien de cause à effet à l'intérieur du triple emboîtement artéfactuel, avec un artéfact à la source du conflit, et un autre qui, en bout de chaîne, est affecté en conséquence. On remarque par ailleurs que les conflits instrumentaux en lien avec les artéfacts pédagogique et didactique sont sur-représentés. Dans la plupart des cas, il s'agit d'une articulation entre ces deux artéfacts. L'artéfact pédagogique est quant à lui présent dans tous les conflits instrumentaux identifiés. Enfin, lorsque l'artéfact numérique est impliqué dans un conflit instrumental, un autre artéfact lui est toujours associé. Toutefois, s'il semble pouvoir constituer un facteur aggravant pour des conflits préexistants – c'est-à-dire en lien avec un autre artéfact –, l'artéfact numérique stricto sensu ne paraît pas avoir généré de conflit instrumental à lui seul.

On voit notamment apparaître une combinatoire conflictuelle – autrement dit une combinaison de conflits instrumentaux dont l'origine et l'impact concernent plusieurs types

d'artéfacts – qui s'enracine dans la qualité linguistique parfois problématique de certaines productions d'écrit, ce qui peut être considéré comme un conflit instrumental d'origine didactique. Les écrits en question peuvent s'avérer particulièrement difficiles à coévaluer, ce qui correspond à un conflit instrumental affectant l'artéfact pédagogique. Or pour rappel, les écrits sont produits sur traitement de texte – théoriquement avec l'aide d'un correcticiel ; toutefois, un tel logiciel peut faire l'objet d'un non-usage, soit parce qu'il n'a pas été activé, soit parce que l'élève ne parvient pas à interpréter les signalements de l'interface graphique, ni les suggestions de correction, et préfère de ce fait les ignorer. Des mésusages peuvent également survenir avec un correcticiel, si les prérequis didactiques précédemment évoqués viennent à manquer : l'interprétation d'une suggestion de correction orthographique automatisée, et le choix éclairé parmi plusieurs formes – aux graphies souvent proches –, peuvent être une source de confusion, singulièrement pour des élèves en difficulté. Enfin, aux dires des élèves, une interface comme le clavier peut être la source (numérique) de conflits instrumentaux affectant l'artéfact didactique, *e.g.* lorsqu'on ne sait pas où se trouve un caractère spécial, un accent circonflexe ou une cédille.

On se trouverait alors en présence d'une combinatoire conflictuelle dont la source consiste en une articulation problématique des artéfacts didactique et technique, et qui rendrait la tâche singulièrement plus difficile pour le pair coévaluateur qui se trouve confronté à de tels écrits. En effet, eu égard à la tâche coévaluative stricto sensu, une surcharge (méta-)cognitive peut survenir, entraînant un retard d'apprentissage – voire un renoncement partiel ou total – vis-à-vis de la tâche de coévaluation (artéfact pédagogique), et de facto, vis-à-vis de l'artéfact didactique également (métacompétences).

4.1.5.2 Degrés de conflictualité

Relativement aux degrés de conflictualité, le Tableau 17 récapitule les conflits instrumentaux observés, ainsi que les artéfacts et degrés de conflictualité afférents.

Conflits instrumentaux	Artéfacts concernés			Degrés de conflictualité		
Problèmes de discernement docimologique	Dida.	Péda.		Retard	Renoncement partiel	
Inhibitions socioaffectives	Dida.	Péda.		Retard		
Inter- vs hétéro-évaluation : vision conflictuelle		Péda.	Tech.	Retard	Renoncement partiel	
Difficulté à formuler un écrit commentatif adapté	Dida.	Péda.		Retard	Renoncement partiel	
Contresens pédagogique à l'égard de la coévaluation	Dida.	Péda.			Renoncement partiel	Renoncement total
Mésusages ou non-usage des interfaces	Dida.	Péda.	Tech.	Retard	Renoncement partiel	
Absence de prérequis didactiques	Dida.	Péda.				Renoncement total
Surabondance d'erreurs orthographiques dans les écrits	Dida.	Péda.	Tech.	Retard	Renoncement partiel	

Tableau 17 – Conflits instrumentaux observés, avec artéfacts et degrés de conflictualité afférents (jaune = retard ; orange = renoncement partiel ; rouge = renoncement total)

On remarque que les conflits instrumentaux les plus problématiques en termes d'apprentissages se rapportent à une articulation entre les artéfacts pédagogique et didactique. Lorsqu'en association avec d'autres artéfacts, l'artéfact numérique est impliqué dans un conflit instrumental, il s'accompagne de retards et ou de renoncements partiels à la tâche d'apprentissage, mais sans aller jusqu'au renoncement total. Enfin, les retards et renoncements

partiels ou totaux peuvent concerner les élèves scripteurs et/ou les élèves coévaluateurs, et se caractériser par une gêne instrumentale affectant l'accès à un ou plusieurs artéfact(s).

Une de ces combinatoires conflictuelles pourrait impliquer en l'occurrence la maîtrise insuffisante d'artéfacts didactiques intervenant en amont de la situation expérimentale d'enseignement/apprentissage (prérequis didactiques). Les résultats concernés sont issus de données autorapportées, ils ont donc une valeur déclarative et ne permettent pas d'accéder au détail des compétences qui sont en jeu. En outre, la terminologie *prérequis didactique* peut recouvrir plusieurs facteurs, *e.g.* intelligibilité des savoirs, capacités cognitives, habiletés procédurales, liens entre les contenus d'apprentissage. Quoi qu'il en soit, ces résultats laissent supposer que l'absence de certains prérequis didactiques gêne ou empêche l'accès à l'artéfact pédagogique et, partant, à l'artéfact didactique cible (retard, renoncement partiel ou renoncement total à l'activité d'apprentissage).

De même, concernant le discernement docimologique des élèves – c'est-à-dire leur capacité métacognitive à apprécier avec justesse le degré de maîtrise d'une compétence –, les difficultés observées peuvent être le signe d'une maîtrise insuffisante des compétences cibles, c'est-à-dire d'une relative absence de prérequis didactiques. Ainsi, un conflit instrumental d'origine didactique est là encore susceptible d'entraîner un retard – voire un renoncement partiel ou total – vis-à-vis de la coévaluation (artéfact pédagogique), mais aussi à l'égard de la composante métacognitive – autrement dit les métacompétences cibles – de l'artéfact didactique.

Singulièrement, nos résultats permettent de mettre au jour des conflits instrumentaux dont la source est d'ordre pédagogique, mais dont les répercussions sont didactiques, et qui affectent un nombre significatif d'élèves. A cet égard, on peut notamment citer les contresens pédagogiques – lorsque l'élève ne perçoit pas la signification et/ou l'utilité de la coévaluation –, la peur de coévaluer un pair de manière injuste (inhibition d'ordre socio-affectif), la difficulté à formuler un écrit commentatif adapté (c'est-à-dire constructif et argumenté) ou une vision conflictuelle opposant inter-évaluation et hétéro-évaluation, fondée sur une vision arbitraire de l'acte évaluatif. Dans tous les cas précités, un conflit instrumental pédagogique – autrement dit un conflit instrumental lié à l'artéfact « coévaluation » – est susceptible d'entraîner un retard ou un renoncement à la tâche d'apprentissage.

De même, on a pu voir l'importance, dans les déclarations d'élèves coévaluateurs, d'un conflit instrumental pédagogique afférent à la difficulté à formuler un écrit commentatif adapté – c'est-à-dire constructif et argumenté. En réception, on constate effectivement que ces écrits commentatifs ne sont pas toujours suffisamment explicites du point de vue des élèves

coévalués. L'attention critique manifestée par les élèves participants à l'égard de cet aspect parfois défaillant du dispositif semble néanmoins indiquer une bonne appropriation des enjeux de ce type spécifique de feedback, à la fois en production et en réception. En effet, leurs propos permettent de supposer qu'il s'agit pour eux, à travers des écrits commentatifs de qualité, d'appuyer à la fois la légitimité de la coévaluation et son acceptabilité, et de permettre à l'élève coévalué d'effectuer des progrès réels, sur la base de retours circonstanciés et justifiés au regard des consignes d'écriture et des compétences cibles. Cette insistance des élèves sur le rôle de l'écrit commentatif – un type de feedback qui, dans le cadre du présent dispositif, relève à la fois des artéfacts pédagogique et numérique –, a donc des implications potentielles significatives relativement à l'artéfact didactique – c'est-à-dire sur les compétences cibles en français langue première (Figure 70).

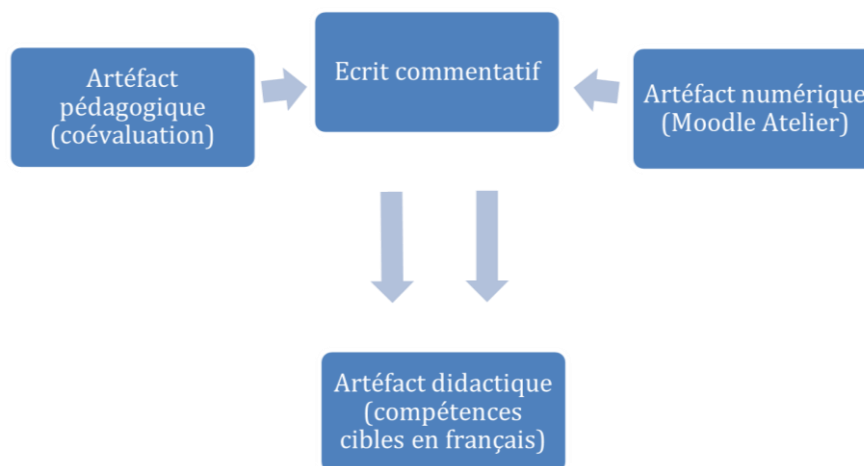


Figure 70 – Rôle pivot de l'écrit commentatif dans le dispositif de coévaluation numérique

4.1.5.3 Synergies régulatrices

Pour finir, on voit apparaître des synergies régulatrices, c'est-à-dire des combinaisons de facteurs de régulation, susceptibles de compenser partiellement ou totalement certains des conflits instrumentaux identifiés. Le Tableau 18 résume ces synergies régulatrices.

	Appétence pour l'artéfact pédagogique	Appui sur expérience acquise	Facilité d'utilisation	Appétence pour l'artéfact technique	Horizontalisation de la posture enseignante	Anonymat : scripteur et inter-évaluateur	Clarté des consignes / « Exemples de travaux »	Présence de prérequis didactiques
Inhibitions socioaffectives								
Inter- vs hétéro-évaluation : vision conflictuelle								
Difficulté à formuler un écrit commentatif adapté								
Contresens pédagogique à l'égard de la coévaluation								
Mésusages ou non-usage des interfaces (clavier, correcticiel...)								
Absence de prérequis didactiques								
Problèmes de discernement docimologique								
Erreurs orthographiques dans les écrits								

Tableau 18 – Conflits instrumentaux et synergies régulatrices (violet = artéfact pédagogique ; orange = artéfact technique ; bleu = artéfact didactique)

Nos résultats montrent par exemple les apports de Moodle Atelier en termes de compréhension des consignes : l'artéfact numérique semble jouer un rôle facilitateur relativement à l'artéfact pédagogique. On peut supposer que ce rôle régulateur relève d'un processus d'instrumentation de l'artéfact numérique, *e.g.* par le truchement de son interface graphique, ou via une fonctionnalité spécifique de l'artéfact numérique, comme les « Exemples de travaux ».

En effet, on constate que la fonctionnalité « Exemples de travaux », spécifique à l'artéfact numérique, constitue à la fois un point d'appui important pour les élèves en termes de compréhension des consignes et un apport significatif du numérique à la situation d'enseignement/apprentissage. De fait, Moodle Atelier permet de scénariser ces « Exemples de travaux » comme une « mini-activité » à part entière : l'élève est invité à évaluer pour s'entraîner, et un retour interactif lui est immédiatement fourni par le logiciel, suivant les directives configurées en amont par l'enseignant. En temps réel, l'apprenant peut ainsi se faire une idée de la justesse des choix évaluatifs qu'il vient d'effectuer, et ajuster son discernement docimologique en conséquence, autrement dit sa capacité métacognitive à apprécier avec justesse le degré de maîtrise d'une compétence par un pair. Eu égard à l'artéfact numérique de coévaluation, il semble que nous nous trouvions ici en présence d'une fonctionnalité régulatrice au sens de (Zeller & Marquet, 2021b), autrement dit une fonctionnalité susceptible de contribuer à la résolution de conflits instrumentaux préalablement identifiés, singulièrement :

- le manque d'expérience des élèves en matière de coévaluation ;
- la difficulté ressentie à évaluer avec pertinence l'écrit d'un pair :
 - en raison de problèmes de discernement docimologique ;
 - du fait d'inhibitions d'ordre socioaffectif.

A l'instar des « Exemples de travaux », l'anonymisation en double aveugle constitue une fonctionnalité spécifique à l'artéfact numérique⁵⁷. En effet, le fait que dans Moodle Atelier on puisse anonymiser facilement – en quelques clics – à la fois les auteurs des productions écrites et les coévaluateurs, constitue une différence significative vis-à-vis d'éventuels dispositifs de coévaluation manuscrits. Cette anonymisation en double aveugle apparaît donc comme un élément déterminant de l'artéfact numérique, qui semble avoir facilité la coévaluation, conformément aux préconisations de (Dochy *et al.*, 1999; Topping, 1998) et, partant, pourrait avoir contribué à l'amélioration, établie *supra*, des compétences en français.

Dans l'ensemble, ces éléments d'étayage pédagonumérique pourraient constituer des points d'appui pour contrebalancer ou réguler certains des conflits instrumentaux que nous avons identifiés *supra*, *e.g.* les mésusages liés aux interfaces de l'artéfact technique ou le « contresens » lié à l'artéfact pédagogique, singulièrement en favorisant le déploiement d'une différenciation pédagogique, notamment à destination des élèves qui peinent à percevoir le sens

⁵⁷ Comme nous l'avons vu *supra*, la mise en œuvre manuscrite de l'anonymisation en double aveugle s'avère difficile voire impossible à pratiquer régulièrement avec une classe de taille moyenne, *i.e.* 30 élèves.

ou l'utilité de la coévaluation. Autrement dit, lorsque l'on considère le triple enchâssement des artefacts didactique, pédagogique et technique dans le contexte d'un EIAH, la composante technique/numérique peut contribuer de façon significative à la genèse instrumentale, en synergie avec la composante pédagogique, facilitant l'accès à l'artéfact didactique.

A ce titre, nos résultats montrent que les élèves se sont sentis particulièrement accompagnés par leurs enseignants. Cela semble particulièrement intéressant compte tenu des résultats de la dernière enquête PISA (OCDE, 2019b), dont nous avons vu *supra* qu'ils révélaient précisément, chez les élèves français, un sentiment d'accompagnement très inférieur à la moyenne des pays de l'OCDE. Or l'artéfact numérique pourrait avoir joué un rôle significatif en ce sens. En effet Moodle Atelier permet à l'enseignant d'automatiser certaines interactions – pédagogiquement scénarisées, et techniquement paramétrées en amont –, ce qui contribue à le rendre davantage disponible pour aider les élèves qui en ont besoin. L'enseignement tend ici à être moins frontal et davantage horizontal : en d'autres termes, structurellement, le dispositif de coévaluation numérique tend à horizontaliser la posture de l'enseignant pour un accompagnement plus individualisé des élèves. En termes de différenciation pédagogique, cela revient à s'appuyer sur un artéfact numérique pour promouvoir un apprentissage autogéré et permettre à l'enseignant de pratiquer une forme ponctuelle de préceptorat (Forget, 2017b, p. 33).

Dans cette perspective, la catachrèse (Rabardel, 1995) suggérée par les élèves participants semble une piste intéressante. Elle consisterait en la mise en œuvre, dans Moodle Atelier par exemple, d'une boucle formative de rétroaction entre coévaluateurs et coévalués, en personnalisant pour chaque apprenant le nombre d'aller-retours nécessaires pour valider une compétence. Ce travail de réécriture, doublement outillé par le traitement de texte (Plane, 1996) d'une part, et par l'artéfact numérique de coévaluation (Bouziane & Ziad, 2018a) d'autre part, constituerait une manière d'opérationnaliser la différenciation pédagogique via des apprentissages autogérés (Forget, 2017b), en appui sur la coopération entre apprenants. Cette suggestion de catachrèse fait donc écho aux notions de collectif apprenant (Lescouarch, 2021) et de boucle évaluative (Connac & Cieutat, 2021; Reynaud, 2022). L'anonymisation en double aveugle – dont nous avons vu *supra* qu'elle constitue une pratique quasi inaccessible en classe entière sans l'instrumentalisation d'un artéfact numérique – jouerait là encore un rôle clé, dans la mesure où elle permettrait de diminuer la peur de noter injustement (inhibitions socioaffectives) ou de pallier d'autres biais sociocognitifs (*e.g.* pression exercée par les pairs).

En somme, nos résultats permettent de valider l'hypothèse HO 4a, c'est-à-dire « La coévaluation numérique génère des conflits instrumentaux affectant les situations d'apprentissage, mais ces derniers peuvent être régulés ».

4.2 Situations d'enseignement (HO 4b)

Le point de vue des enseignantes partenaires de la présente recherche constitue un contrepoint indispensable à celui des élèves participants, à plus forte raison dans la mesure où il s'agit d'une recherche-intervention à caractère collaboratif. En appliquant les principes méthodologiques détaillés *supra*, six entretiens semidirigés et un focus groupe ont été mis en œuvre. Les enregistrements audios de ces temps d'échange⁵⁸ sont accessibles en Annexes 6 et 7. Les résultats sont présentés ici par méthode de recueil et par ordre chronologique.

4.2.1 Entretiens semidirigés

4.2.1.1 Enseignante A (T2)

Avec l'enseignante A, un bref entretien semidirigé a été réalisé à T2, immédiatement après la première expérimentation avec l'échantillon A. Dans la chronologie de la présente étude, il s'agit de la toute première mise en œuvre du protocole expérimental. Il en est ressorti un élément saillant, à savoir le sentiment que le dispositif ne semble pas permettre à l'enseignante de mesurer si les élèves ont effectivement pu progresser dans leurs apprentissages, à la différence de ce qui se passe « quand on est en classe normalement ».

4.2.1.2 Enseignante C (T2)

L'entretien semidirigé réalisé à T2 avec l'enseignante C porte essentiellement sur les composantes didactique et pédagogique du dispositif mis en œuvre dans le cadre de cette recherche collaborative. Précisément, l'enseignante C met en avant l'importance de la lecture : en amont de l'activité de production écrite, c'est la lecture d'un « grand texte » qui sert de déclencheur. Par ailleurs, selon l'enseignante C, dans la grille de coévaluation co-construite avec les élèves selon l'approche par compétences, « la cohérence avec le texte source [est] extrêmement présente », ce qui constitue « un choix appuyé par les élèves : cohérence verbale,

⁵⁸ A l'exception de l'entretien semidirigé mené à T2 avec l'enseignante C, réalisé en présentiel, sans enregistrement audio, avec prise de notes de la part du chercheur.

cohérence sémantique, cohérence stylistique ». Une captation d'écran représentant la grille de coévaluation en question (classe C, 5^e, T1-T2) est consultable en Annexes 2.1.2.2. Parmi les compétences ciblées, on voit notamment apparaître une contrainte d'ordre lexical : « Le sens de l'ouïe est mis en valeur » et une contrainte stylistique ou syntaxique : « Le texte comporte au moins une comparaison » – sachant qu'ici la figure de style « comparaison » ou le degré de l'adjectif sont tous deux considérés comme valables. D'après les dires de l'enseignante C, les nuances docimologiques présentées dans le Tableau 19 n'ont pas été suffisamment explicitées en amont de l'activité ; cependant, à en juger par les questions qu'ils ont posées en classe, une minorité d'élèves semblent se les être appropriées.

Niveau de maîtrise (échelle de Likert à quatre points)	Descriptif
« Insuffisant »	L'élève n'a même pas essayé.
« Fragile »	L'élève a essayé mais maladroitement (la réussite n'a pas été au rendez-vous).
« Satisfaisant »	Une comparaison réussie.
« Très satisfaisant »	Plus d'une comparaison réussie (ou une particulièrement originale).

Tableau 19 – Niveaux de maîtrise et descriptifs issus de la réflexion docimologique menée par l'enseignante C avec ses élèves de 5^e à T1-T2 pour l'item « Comparaison » de la grille de coévaluation afférente

4.2.1.3 Enseignante B (T4)

L'entretien semidirigé avec l'enseignante B révèle que la principale problématique pour elle avec cette expérimentation est celle du temps que cela prend : à la fois pour la co-construction de la séance et pour le paramétrage de Moodle Atelier, même si ce dernier a essentiellement été pris en charge par le chercheur. L'expérimentation avec le dispositif pédagonumérique de coévaluation a également pour effet une objectivation accrue de l'évaluation pour l'enseignante B, avec en particulier une formulation plus rigoureuse et plus précise des critères. Selon ses dires, la dématérialisation liée à l'outil numérique rend par ailleurs les productions écrites des élèves quasi anonymes, ce qui entraîne pour elle un surcroît d'objectivité. Une anonymisation totale n'est cependant pas souhaitable selon elle, dans la mesure où la différenciation pédagogique passe également par la capacité à adapter son geste évaluatif au profil de l'élève, à l'historique de ses progrès ou à ses efforts – à fortiori dans le cadre d'une évaluation formative. Outre ce quasi-anonymat des travaux d'élèves à ses yeux,

l'enseignante B évoque une rationalisation de son processus évaluatif, notamment parce qu'elle se trouve délivrée de certains facteurs parasites comme « la lisibilité de la copie », grâce à la dématérialisation des productions écrites, mais aussi parce que l'interface de Moodle Atelier permet une simplification du geste évaluatif stricto sensu, directement translaté dans une grille de coévaluation ergonomique. L'enseignante B déclare en outre que sa participation à la présente recherche renforce chez elle une démarche réflexive, déjà engagée par ailleurs, concernant le choix des modalités évaluatives et des compétences ciblées, en fonction des objectifs de telle ou telle situation d'apprentissage. Selon ses dires, cette réflexion s'accompagne chez elle d'une prise de conscience de la nécessité de restreindre le nombre de compétences évaluées, pour que cela soit plus lisible pour les élèves : « Trop de grille tue la grille ». L'enseignante B déclare également qu'elle envisage une restriction du nombre de critères dans les grilles de ses propres hétéro-évaluations sommatives. Pour elle, cette expérimentation avec Moodle Atelier fonctionne en outre comme un point d'entrée vers Moodle en général et l'amène à insérer dans un cours Moodle l'ensemble de sa séquence d'enseignement.

Concernant ses élèves, l'enseignante B estime que la co-construction de la grille avec eux a été fructueuse dans la mesure où elle a permis de favoriser l'appropriation et l'acceptation des critères par eux. Des gains en termes d'apprentissages sont de surcroît signalés, essentiellement sur un plan métacognitif : une meilleure compréhension des attendus, des critères de réussite ; concrètement, en jouant en quelque sorte le rôle du professeur. Un conflit instrumental est cependant signalé concernant l'artéfact didactique. Selon l'enseignante B, les élèves les plus en difficulté n'ont pas pu entrer dans l'activité métacognitive qu'est la coévaluation en raison d'un niveau de compétence trop faible. L'enseignante partenaire précise que, pour ces élèves, des problématiques de motivation et d'autonomie viennent se surajouter. Selon elle, y compris avec une activité comme celle-ci, qui a recours aux pédagogies du numérique pour proposer quelque chose de différent, on n'obtient peu d'engagement, d'adhésion ou de prise au sérieux de l'activité. Elle estime que les élèves qui sont déjà en réussite au préalable continuent à progresser à la faveur du dispositif ; ainsi, les écarts entre élèves, en termes de maîtrise des compétences linguistiques, continuent de se creuser.

Pour finir, l'enseignante B aimerait pouvoir systématiser ce type d'activité, mais en distanciel asynchrone, sous forme de devoirs à faire régulièrement à la maison. Cela présenterait l'avantage de ne pas empiéter sur les heures de cours, tout en permettant de rendre ce type d'activité habituel. Cela entraînerait les élèves à la pratique de l'inter- et de la coévaluation, et cela pourrait donner encore plus de sens à ces modalités évaluatives. Un tel

fonctionnement suppose cependant que les élèves fassent leurs devoirs à la maison et accomplissent plus ou moins seuls les démarches liées à la participation à un atelier de coévaluation sur Moodle : production écrite, téléversement dans la plateforme logicielle, évaluation des pairs ; or l'enseignante B n'est pas certaine que des élèves de cycle 4 disposent du degré d'autonomie suffisant, en particulier certains élèves très peu motivés voire en situation de quasi-décrochage scolaire.

4.2.1.4 Professeure-documentaliste du collège A (T4)

A T1-T2, le nombre d'ordinateurs fonctionnels en salle informatique s'est avéré insuffisant ; en conséquence, nous avons demandé à la professeure-documentaliste si elle voulait bien accueillir quelques élèves pour qu'ils puissent réaliser l'activité de coévaluation dans Moodle Atelier sur des postes informatiques du centre de documentation et d'information (CDI) du collège, sous sa supervision. La professeure-documentaliste a ainsi accepté d'accueillir trois élèves, qui ont travaillé individuellement, chacun sur son poste informatique. En théorie, conformément au protocole expérimental co-construit avec les enseignantes partenaires, les trois élèves concernés n'étaient pas censés travailler en groupe, si ce n'est en coopérant avec le reste de la classe par l'entremise de l'activité de coévaluation sur Moodle Atelier, c'est-à-dire de manière synchrone et présentielle, par ordinateur interposé.

Or lors de l'entretien semidirigé, la professeure-documentaliste déclare qu'à plusieurs reprises les trois élèves qu'elle a accueillis au CDI se sont concertés pour rechercher la plus grande justesse possible dans leur évaluation des niveaux de maîtrise à attribuer à telle ou telle production d'écrit, pour tel ou tel critère. A titre d'exemple, elle témoigne du fait que, parmi eux, une élève avait du mal à se décider quant au niveau de maîtrise à choisir pour évaluer une des compétences cibles, à savoir la première dans le formulaire de coévaluation ci-après (Figure 71).

Formulaire d'évaluation ▼

Critères	Niveaux			
- le texte contient des informations : date de création, acteur, décor, thème, intrigue, metteur en scène, dramaturge... > décrire ce qu'on voit > employer le vocabulaire spécifique au théâtre	<input type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input checked="" type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
- formuler un jugement sur ce qu'on voit (adjectif, verbe d'opinion...)	<input checked="" type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
- forme du texte : paragraphe, texte organisé, connecteurs...	<input checked="" type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
- la critique doit montrer qu'on a bien compris le texte.	<input checked="" type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant

Figure 71 – Grille de coévaluation pour l'activité de coévaluation à T1-T2 avec l'échantillon A'

Au dire de la professeure-documentaliste, cette élève a spontanément demandé à ses deux camarades ce qu'ils pensaient du fait qu'en l'occurrence le pair évalué n'avait « pas parlé des acteurs ». Ses deux camarades se sont alors accordés à dire que dans ce cas, on ne pouvait pas attribuer à la production écrite concernée le niveau de maîtrise « Très satisfaisant ». Cet exemple de « débat docimologique » paraît d'autant plus intéressant que, d'après la professeure-documentaliste, les trois élèves en question « ne sont pas forcément très copains, à la base ». Il semble que « le fait qu'ils soient seulement tous les trois [au CDI], dans un lieu où il n'y avait personne d'autre » a été particulièrement propice à ces échanges. La professeure-documentaliste fait par ailleurs remarquer que ces échanges étaient très horizontaux, sans que l'un des trois ne cherche à prendre l'ascendant dans les discussions : « ils faisaient ça de façon très positive, à égalité ». Enfin, signalons qu'un des élèves – désigné par la professeure-documentaliste comme étant doté des compétences scolaires les plus solides parmi les trois élèves présents ce jour-là – s'est réjoui de « prendre la place du prof », notamment concernant la rédaction d'écrits commentatifs à destination des pairs.

4.2.1.5 Enseignante C (T6)

Lors de l'entretien semidirigé réalisé à T6, l'enseignante C met tout d'abord en avant le fait que la présente recherche collaborative a constitué un projet fédérateur, une ouverture de la classe sur le monde et une stratégie de valorisation pour les productions écrites, qui a incité chaque apprenant à donner le meilleur de lui-même. L'enseignante C précise que les élèves eux-mêmes sont porteurs du projet, ils sont investis – en quelque sorte comme des

ambassadeurs spontanés. En effet, ils sont enclins à y sensibiliser des élèves d'autres classes, qui à leur tour peuvent en parler à leurs enseignants, parce qu'ils aimeraient expérimenter la coévaluation numérique. A ce titre, l'enseignante C signale d'ailleurs que les élèves peuvent être force de proposition en matière d'artéfacts pédagogiques ou d'artéfacts techniques/numériques, et contribuer ainsi à l'essaimage de bonnes pratiques ou de pratiques innovantes au sein de l'établissement, c'est-à-dire d'enseignant à enseignant, par l'entremise des élèves. L'enseignante C cite l'exemple des diaporamas : ce sont des élèves qui lui ont suggéré de modifier ses schèmes d'enseignement pour intégrer le diaporama à ses pratiques ; et c'est suite à leurs suggestions, et grâce à leur aide ainsi qu'à celle de collègues, qu'elle a pu s'approprier – c'est-à-dire instrumenter – des artéfacts techniques comme Microsoft Powerpoint, Google Slides ou LibreOffice Impress, afin de réaliser des diaporamas qui, selon ses dires, ont très bien fonctionné auprès des élèves. L'enseignante C estime que le dispositif expérimental de coévaluation pourrait bénéficier d'un mode de diffusion comparable.

Evoquant le cas de certains élèves en grande difficulté, l'échange porte alors sur un cas particulier : celui d'un de ces élèves, qui faisait auparavant preuve d'une grande inhibition, avec un blocage face aux activités de production écrite, notamment en raison d'une maîtrise très fragile de l'orthographe lexicale et de la morphosyntaxe – ce qui le plaçait en situation d'échec répété. L'enseignante C signale alors qu'elle a été agréablement surprise concernant les progrès accomplis par cet élève dans le cadre de l'expérimentation. D'une part, l'élève en question s'est investi dans les ateliers d'écriture avec Moodle Atelier, dès les phases de co-construction de la grille de coévaluation, puis de production écrite stricto sensu. Selon l'enseignante C, la grille de coévaluation pourrait avoir constitué un étayage pédagogique important pour cet élève, probablement parce que, par l'entremise de l'artéfact pédagogique, cette grille s'inscrivait dans un cadre coopératif, en l'occurrence : « la façon dont il allait [coévaluer] allait avoir un impact sur ses camarades ». Surtout, lors de la phase de coévaluation proprement dite, cet élève a manifesté « une attitude d'engagement très fort » vis-à-vis de l'activité, en allant spontanément chercher son cahier de français pour vérifier un point de langue susceptible d'éclairer son jugement, eu égard au geste évaluatif qu'il était en train de poser sur la production écrite d'un de ses pairs. Or l'enseignante C signale que tous les élèves disposaient déjà, pour ce faire, de fiches-outils ou fiches-ressources – à la fois au format papier et sous forme dématérialisée, incluses dans le cours en ligne, à même la plateforme Moodle – très facilement consultables. Cet élève a donc accompli la démarche d'aller chercher une ressource supplémentaire pour nourrir sa réflexion et mieux accomplir la tâche, ce qui

manifeste un « savoir-être », relatif à l'autorégulation des apprentissages, très peu présent chez lui jusqu'ici d'après l'enseignante C.

L'enseignante partenaire est également amenée à faire part de son ressenti relativement aux éléments de cadrage induits par le dispositif pédagonumérique – c'est-à-dire avec l'artéfact didactique en ligne de mire, mais singulièrement par la convergence, pour la genèse instrumentale du dispositif, des artéfacts numérique et pédagogique en particulier. Ces éléments de cadrage concernaient notamment l'approche par compétence, le « formulaire d'évaluation » commun pour l'hétéro-évaluation de l'enseignant et l'inter-évaluation entre élèves, ainsi que l'obtention d'une note supplémentaire, inhabituelle pour les élèves et spécifique à l'artéfact technique – en l'occurrence la note de coévaluateur. L'enseignante C déclare à cet égard que la co-construction des critères de réussite constituait déjà une pratique courante pour ses élèves et elle. Elle signale toutefois que l'expérimentation dont il est question ici lui a permis d'approfondir sa réflexion vis-à-vis de ses pratiques évaluatives, notamment pour les affiner, les rationaliser et les rendre plus explicites. Elle évoque en particulier la prise de conscience de la nécessité d'un accroissement de la « finesse docimologique » : « cette idée, maintenant, qu'il faille décider à quel moment on attribue telle ou telle note, en fonction de la qualité de réussite sur un critère » – là où le processus évaluatif était auparavant d'avantage intuitif et moins rationnel/objectif, donc moins transparent pour les élèves.

En outre, elle signale un événement remarquable survenu avec deux élèves en particulier, lors de la dernière séance dévolue à la production écrite. Certains élèves avaient déjà terminé, d'autres étaient en retard, pour différentes raisons. Ceux qui avaient déjà terminé avaient la possibilité d'effectuer une production écrite subsidiaire, « sur un thème assez libre ». Parmi eux, deux élèves se sont assis côte à côte et ont réalisé chacun leur production écrite – en s'entraînant, même si le travail était de nature très personnelle – et ont commencé, sans s'en rendre compte, à se fixer des contraintes pour leurs productions écrites. L'enseignante a accompagné ce mouvement en leur proposant de se fixer d'autres contraintes d'ordre linguistique pour améliorer la qualité de leur production écrite. Selon l'enseignante C, on serait passé d'une contrainte extrinsèque – qui peut souvent être vécue comme un frein à la réussite des apprentissages, à fortiori par des élèves en situation d'échec scolaire ou confrontés à des difficultés socioculturelles, d'ordre familial ou personnel, qui se traduisent par une relation abîmée à l'adulte et à l'institution scolaire en général – à une contrainte intrinsèque.

Un autre aspect, connexe, évoqué lors de l'entretien avec l'enseignante C est celui du choix et de la formulation des consignes, critères et/ou compétences cibles dans le formulaire d'évaluation de Moodle Atelier, *e.g.* pour les rendre si possible moins artificiels, selon une

approche renouvelée de la didactique de la langue. Prenant le cas de Marie⁵⁹, élève dont les compétences en français langue première sont jugées très solides par l'enseignante C, elle fait part de la difficulté de cette dernière à s'appropriier des consignes d'écriture vécues comme des contraintes extrinsèques. Il s'agissait en l'occurrence de faire une métaphore. Selon l'enseignante partenaire, Marie a incorporé à son texte une métaphore un peu « forcée », qui n'était pas en adéquation avec le caractère du personnage concerné. L'enseignante C estime donc qu'il serait intéressant d'interroger Marie pour lui demander dans quelle mesure la découverte des productions écrites de ses pairs lui a permis d'envisager la consigne, non plus comme une contrainte extrinsèque, mais comme une contrainte intrinsèque, intériorisée, acceptée car source de créativité. Or dans son retour écrit sur expérience, Marie a écrit :

« J'ai beaucoup aimé le fait de pouvoir découvrir comment nos camarades ont imaginés leurs textes/suites de textes, car nous n'avons pas tous les mêmes idées et c'est génial de pouvoir les partager. »

A la lumière de ce retour écrit sur expérience, consulté en temps réel pendant l'entretien semidirigé, l'enseignante partenaire estime que Marie a bénéficié de l'emboîtement des artefacts pédagogique et numérique, pour un accès facilité à l'artéfact didactique.

Pour l'enseignante C, l'articulation entre étude de la langue et production d'écrit (artéfact didactique) se trouve également renouvelée dans ses pratiques. Ses schèmes d'enseignement évoluent à la faveur d'un phénomène d'instrumentation, « de l'artéfact vers le sujet ». En l'occurrence, par l'entremise des artefacts technique et pédagogique, l'enseignante C est désormais amenée à travailler les compétences cibles de manière plus organique avec ses élèves, c'est-à-dire en s'appuyant non seulement sur des exemples tirés de l'œuvre intégrale étudiée en classe, mais également sur des extraits de productions écrites d'élèves, illustrant des points de langues à l'étude – voire suscitant l'approfondissement de certaines notions, *e.g.* des degrés de l'adjectif à la métaphore et à la métaphore filée, en passant par la comparaison comme figure de style. Au dire de l'enseignante interrogée, ces schèmes d'enseignement étaient déjà en place auparavant, avec « des structures de cours entières qui [étaient] le fruit d'une réflexion avec les élèves » ; ils ont toutefois été significativement favorisés, en amont, par la circulation anonymisée des productions d'écrits entre les élèves, par l'entremise de

⁵⁹ Comme indiqué *supra*, les prénoms d'élèves ont été changés pour préserver leur anonymat, conformément au formulaire de consentement libre et éclairé signé par leurs parents et conformément au règlement général de protection des données (RGPD).

l'artéfact pédagogique « coévaluation » – lui-même facilité par l'artéfact numérique « Moodle Atelier ».

Pour finir, l'enseignante C signale une relative absence de conflit instrumental en lien avec l'artéfact technique, notamment pour se connecter aux ordinateurs, à l'environnement numérique de travail et accéder à Moodle Atelier : la quasi-totalité des élèves de l'échantillon C connaissent en effet leurs codes ENT, et tous ont *in fine* réussi à se connecter rapidement. L'enseignante partenaire estime que le fait qu'il s'agisse d'une « classe coopérative » joue un rôle important à cet égard. En effet, une « classe coopérative » est une classe pour laquelle les pédagogies coopératives font l'objet d'une pratique généralisée de la part de l'ensemble de l'équipe enseignante, dans le cadre d'un projet d'établissement, accompagné par le rectorat d'académie au moyen d'un programme ad hoc de formation continue⁶⁰. L'enseignante C précise que cette relative absence de conflit instrumental technique a été particulièrement prégnante dans le cadre de l'expérimentation menée avec le dispositif péda-numérique de coévaluation. En outre, concernant les différentes opérations nécessaires à la manipulation de l'artéfact numérique et à son appropriation par les élèves de l'échantillon C – du téléversement des productions écrites dans la plateforme, aux opérations de coévaluation et de rédaction d'un écrit commentatif pour les pairs –, l'enseignante C estime qu'on a également pu constater une relative absence de conflit instrumental.

4.2.1.6 Enseignante B (année N+1)

Pour rappel, lors de l'année N+1 (2022-2023), avec une nouvelle classe de 3^e – distincte de l'échantillon B avec lequel nous avons expérimenté lors de l'année N –, en autonomie (c'est-à-dire sans aucune intervention du chercheur, en tout cas en amont de la mise en œuvre) et à deux reprises, l'enseignante B s'est réapproprié l'artéfact pédagogique « coévaluation » par l'entremise de l'artéfact numérique Moodle Atelier. Les détails didactiques et pédagogiques de ces réappropriation quasi autonome hors protocole sont fournis *supra*. A posteriori, un entretien semidirigé a permis de faire le point sur ces séances menées en autonomie par l'enseignante. En effet, ces séances nous semblent particulièrement intéressantes, d'une part dans la mesure

⁶⁰ Ces formations territoriales ont pour objectif d'accompagner la mise en œuvre des pratiques coopératives localement, « avec l'aide de la recherche et de l'expérience [des] pairs [pour] les mettre en œuvre au sein de [l'équipe] éducative » (cf. Plan académique de formation 2021-2022 de l'Académie de Montpellier, p. 63, lien URL consulté le 11/05/2023 à l'adresse : <https://pedagogie.ac-montpellier.fr/sites/default/files/2021-08/080721%20Fiches-formations-cahier-final-OL.pdf>).

où elles nous informent sur la manière dont une des trois enseignantes partenaires de la présente recherche collaborative s'est approprié le dispositif – ce qui en soi constitue un prolongement spontané du projet de recherche ; et d'autre part pour faire l'inventaire des difficultés rencontrées (conflits instrumentaux). Comme pour l'ensemble des entretiens semidirigés de la présente section, le principal cadre théorique mobilisé pour l'analyse est donc celui de la genèse instrumentale, singulièrement eu égard au triple enchâssement des artefacts didactique, pédagogique et technique en environnement informatisé pour l'apprentissage humain.

Ces deux réappropriations autonomes du dispositif ont lieu en distanciel asynchrone d'une part, et selon un scénario pédagogique hybride, articulant distanciel et présentiel d'autre part. L'enseignante B dit avoir rencontré des difficultés et procédé par tâtonnements. Elle déclare n'être parvenue que partiellement à ses fins, d'où le recours à l'aide du chercheur à posteriori, de façon distancielle et asynchrone, pour finaliser le paramétrage de Moodle Atelier et permettre l'achèvement de l'activité conformément aux objectifs de son scénario pédagogique. Dans les deux cas, les difficultés rencontrées, détaillées ci-après, n'ont pas empêché l'enseignante B de déployer son scénario pédagogique, ni de susciter la mise en activité des élèves, ce qui inclut le recueil de leurs productions écrites – c'est-à-dire leur téléversement dans Moodle Atelier –, ainsi que l'activité de coévaluation proprement dite.

Cette fois, le formulaire d'évaluation fourni dans Moodle Atelier comportait délibérément un nombre restreint de critères – c'est-à-dire six compétences : à ce titre, l'enseignante B a déclaré s'être fondée sur les expérimentations passées, lors de l'année N. En effet, se souvenant de son constat à T4 en année N – relayé *supra* –, à savoir que « Trop de grille tue la grille », elle a procédé à une régulation instrumentale. Le formulaire d'évaluation construit pour cette activité est consultable en Annexes 2.4.1 (Figure 104). On peut en outre constater que l'enseignante B a proposé ici une innovation intéressante par rapport au protocole co-construit en année N : il ne s'agit plus ici, pour l'ensemble des compétences évaluées, d'une seule et même échelle de Likert à quatre points – allant de « Insuffisant » à « Très satisfaisant » –, mais de descriptifs ad hoc, adaptés à chaque critère, avec 2, 3 ou 4 niveaux de maîtrise selon les cas. Aux dires de l'enseignante partenaire, cet étayage métacognitif pourrait avoir aidé les élèves, d'une part à mieux s'approprier les critères d'évaluation pour leurs propres productions d'écrit, et d'autre part à mieux évaluer leurs pairs.

La modalité distancielle asynchrone a posé des problèmes aigus de participation : moins de la moitié des élèves ont déposé leur travail dans Moodle Atelier lors de cette première activité. Moins d'un cinquième de la classe a évalué des pairs, et seulement 15% de l'effectif a procédé à l'auto-évaluation. Selon l'enseignante B, cette faible participation est probablement

liée aux problématiques motivationnelles (échec scolaire, décrochage) d'un nombre significatif d'élèves dans cette classe⁶¹, à fortiori concernant les devoirs à la maison : en effet, en dehors du dispositif « Devoirs faits », un certain nombre d'élèves de l'échantillon B' ne bénéficiaient pas nécessairement des conditions favorables pour un prolongement efficace des apprentissages en dehors du temps scolaire. Ceci dit, les difficultés de paramétrage rencontrées par l'enseignante B pour instrumentaliser Moodle Atelier – c'est-à-dire pour accommoder l'artéfact technique aux schèmes d'utilisation liés à son scénario d'auto- et de coévaluation (artéfact pédagogique) – ont temporairement empêché les élèves de réaliser l'activité. La problématique rencontrée ici s'apparente donc à un conflit instrumental technique ayant retardé les apprentissages des élèves. A la demande de l'enseignante B, l'intervention du chercheur a posteriori a permis de déterminer l'origine du blocage : un défaut de paramétrage. En l'occurrence, l'enseignante B avait activé par erreur, dans les paramètres de Moodle Atelier, une fonctionnalité entraînant le regroupement aléatoire des élèves en petits groupes de travail, ce qui bloquait l'attribution aléatoire automatisée des travaux d'élèves à des pairs pour la phase de coévaluation. Par conséquent, les élèves attendaient, chez eux, que des travaux leur soient attribués afin de pouvoir effectuer les auto-évaluations et les inter-évaluations prescrites. Or il s'agissait de devoirs de vacances. Le problème n'a donc véritablement pu être solutionné qu'à la reprise des cours ; entretemps, un nombre significatif d'élèves avaient été démobilisés par l'attente, notamment ceux dont les compétences sont les plus fragiles et qui ont le plus de difficulté à s'engager – ou à demeurer engagés – dans les apprentissages. Le problème a donc pu être solutionné, mais partiellement seulement, dans la mesure où certains élèves, concernés par les problématiques d'engagement évoqués plus haut, n'ont jamais remis leur travail ni/ou évalué leurs pairs. Ainsi, pour un tiers de la classe, le conflit instrumental n'a pas résulté en un simple retard, mais en un renoncement partiel à l'activité proposée – e.g. remise du travail sans auto- et/ou inter-évaluation, ou inter-évaluation sans remise du travail ; tandis que pour plus de la moitié des élèves, le conflit instrumental a entraîné un renoncement total à l'activité proposée : pas de travail remis, pas d'auto-évaluation, ni d'inter-évaluation.

Au dire de l'enseignante B, ce faible taux de participation – c'est-à-dire ce taux élevé de renoncement partiel ou total à l'activité d'apprentissage – est cependant à nuancer compte tenu du profil de la classe concernée. *In fine*, l'enseignante considère que l'activité demeure une réussite, en particulier compte tenu des très bons résultats obtenus en production écrite, qui

⁶¹ Pour rappel, l'établissement B est situé en réseau d'éducation prioritaire (REP).

témoignent du fait que les objectifs ont globalement été atteints en termes d'apprentissages. Elle met notamment cette réussite sur le compte de l'instrumentalisation de l'artéfact numérique Moodle Atelier, qui a contribué dans une certaine mesure à ce que l'activité proposée sorte de l'ordinaire et soit engageante, attractive, interactive, voire ludique.

Concernant la seconde itération, le jour où la coévaluation numérique a eu lieu, en présentiel et en salle informatique, la moitié des élèves avaient fait le travail préparatoire demandé. Par ailleurs, du temps a dû être pris pour régler divers problèmes d'ordinateur, *e.g.* codes d'accès à l'ENT manquants. Certains élèves en ont profité pour visionner rapidement les bandes-annonces et copier-coller des contenus Wikipédia dans leurs fiches signalétiques. L'enseignante s'en est aperçue : elle a pu partiellement vérifier cela et supprimer certains de ces contenus plagiés en temps réel, pour amener les élèves concernés à se mettre réellement en activité. Les sept critères de coévaluation retenus par l'enseignante B pour cette seconde production écrite sont consultables en Annexes 2.4.2 (Figure 105).

Concernant le regroupement par erreur en petits groupes de travail, le conflit instrumental a pu être réglé cette fois-ci, mais au prix d'un retard d'enseignement (allongement du temps de préparation). Par ailleurs, les 3 opérations de paramétrage avancé qui permettaient de détourner l'usage initialement prévu (évaluation entre pairs) pour aboutir à la catachrèse « coévaluation » faisaient également défaut. Cela a pu être réglé car l'enseignante B a sollicité l'aide du chercheur à temps (c'est-à-dire avant l'attribution des travaux) : le renoncement partiel à l'activité d'enseignement – en l'occurrence à l'anonymisation et à la coévaluation – a ainsi pu être évité. Dans le cas contraire, l'instrumentalisation aurait été obstruée : intégration impossible de l'hétéro- et de l'inter-évaluation, donc impossibilité de pratiquer la coévaluation – et instrumentation forcée, puisque les schèmes d'utilisations initialement associés à l'artéfact technique auraient contraint l'enseignante B à mettre en œuvre une activité d'inter-évaluation exclusivement.

De fait, le conflit instrumental lié à l'environnement numérique pour cette situation d'enseignement/apprentissage – *e.g.* problèmes en salle informatique avec certains ordinateurs, codes ENT manquants pour certains élèves –, s'est combiné à la problématique d'engagement de certains élèves (devoirs non faits). Surtout, l'enseignante B suppose qu'un contresens didactique est probablement survenu pour certains de ces élèves qui n'avaient pas fait chez eux le travail préparatoire demandé : « Ils ont cru que je les interrogeais sur un film, et ont cru être plus malins que moi en allant chercher les infos sur un site ». On peut ainsi considérer que ce contresens concernant l'artéfact didactique a entraîné chez ces élèves un détournement de l'artéfact pédagogique au moyen de l'artéfact numérique – en l'utilisant pour aller chercher les

informations sur internet plutôt que pour effectuer l'activité de production écrite puis de coévaluation initialement prévue. Ce conflit instrumental a pu être partiellement régulé par l'enseignante, dans la mesure où elle a pu identifier certains de ces plagiats et amener les élèves à retravailler leur écrit pour qu'il devienne une production plus personnelle. On peut néanmoins considérer qu'il en a résulté pour ces élèves comme pour l'enseignante un conflit instrumental entraînant un retard et le renoncement partiel à la situation d'enseignement/apprentissage.

De surcroît, pour certains élèves qui ont dû évaluer, sans le savoir, le résultat du plagiat de certains pairs – *e.g.* copié-collé issu de Wikipédia –, l'accès à la composante pédagogique du dispositif a été gêné. Le renoncement à l'activité d'apprentissage n'est cependant que partiel : pour les élèves plagiaires d'une part, dans la mesure où, au dire de l'enseignante B, « ils ont quand même produit quelque chose » – surtout quand on tient compte du fait qu'il s'agit d'une classe globalement difficile, avec de nombreux élèves en situation d'échec scolaire et qui sont très peu engagés dans les apprentissages en classe. D'autre part, on peut parler de renoncement partiel pour les élèves évaluateurs également, y compris lorsqu'ils ont évalué des écrits de pairs plagiaires, au sens où la coévaluation (artéfact pédagogique) au moyen de Moodle Atelier (artéfact technique) leur permet malgré tout d'aiguiser leurs métacompétences, relativement aux critères de réussite pour la réponse écrite à une question de compréhension type DNB (artéfact didactique).

Lors de l'entretien, un autre ajustement est envisagé par l'enseignante B pour pallier ce type de conflit instrumental à l'avenir : un scénario en présentiel synchrone – auquel cas tous les élèves pourraient librement faire des recherches sur internet pour réaliser l'écrit de travail – mais uniquement sous forme de prise de notes, afin de favoriser l'appropriation des contenus et d'empêcher le plagiat mécanique du copié-collé en quelques clics. Alternativement, l'enseignante B envisage un scénario toujours en distanciel asynchrone pour la phase préparatoire, mais en autorisant uniquement les élèves qui ont fait le travail préparatoire à passer à la phase de production écrite – dans le cas contraire le travail serait considéré comme non fait et une autre activité leur serait proposée.

L'enseignante partenaire déclare par ailleurs que, grâce aux expérimentations menées conjointement lors de l'année N dans le cadre de la présente recherche collaborative – et bien qu'elle eût déjà une connaissance préalable de Moodle, en lien avec la formation continue des enseignants (plateforme [M@gistère](#)) – elle a développé une utilisation de plus en plus fréquente de Moodle, en instrumentalisant cet artéfact numérique avec des schèmes d'enseignement qui lui sont propres, et qui visent à en faire « un outil bienveillant, qui leur rend des services ». En effet, dans le but d'aider ses élèves à progresser – en particulier ceux dont les compétences sont

les plus fragiles –, elle entend leur fournir ainsi un espace numérique de travail unique et motivant parce que gratifiant, dans la mesure où l'ensemble du cours, des exercices (d'appropriation, d'entraînement ou de remédiation) et des évaluations (diagnostiques, formatives ou sommatives) y sont structurés, faciles d'accès ; alors qu'un nombre significatif d'élèves de l'échantillon B' peuvent habituellement être enclins à égarer leurs cahiers et autres contenus liés au cours, ce qui nuit aux apprentissages – c'est-à-dire empêche d'ordinaire l'accès à l'artéfact didactique. Ainsi, cette instrumentalisation de la plateforme Moodle par l'enseignante B – conséquence directe, selon ses dires, de la catachrèse « coévaluation » appliquée à l'artéfact Moodle Atelier lors des coexpérimentations en année N avec le chercheur-intervenant – constitue une stratégie de régulation de ce conflit instrumental pédagogique qu'elle rencontrait par ailleurs, que le recours à l'artéfact numérique Moodle en général – autrement dit à l'ensemble des activités qu'il est possible d'y déployer en plus de la seule activité « Moodle Atelier » – semble lui avoir permis de pallier. Ce constat fait écho à ses déclarations lors du premier entretien semidirigé mené à T4, concernant Moodle Atelier en tant que « point d'entrée » vers une utilisation accrue de Moodle en général avec ses élèves. A cet égard, s'agissant des élèves plagiaires, elle a estimé que cette activité a pu être porteuse, au sens où cela a probablement contribué à ancrer l'utilisation de Moodle dans leurs habitudes, en associant potentiellement à cette fréquentation de l'artéfact numérique une connotation positive pour eux, en termes de bénéfices pour leurs apprentissages – et pour leur moyenne trimestrielle.

Pour finir, l'enseignante partenaire souligne qu'elle entend continuer à utiliser non seulement Moodle en général, mais son activité Moodle Atelier en particulier, selon des schèmes d'utilisation similaires – et donc en effectuant les mêmes opérations d'instrumentalisation pour y mettre en œuvre la catachrèse « coévaluation ». Elle a précisé qu'elle envisage de ne pas laisser s'écouler trop de temps entre deux itérations, compte tenu de la relative complexité des opérations de paramétrage avancé, nécessaires à la mise en œuvre de ladite catachrèse. Là encore, il s'agit d'un ajustement de ses schèmes d'utilisation – en l'occurrence une mise en œuvre plus fréquente du dispositif pour mieux se l'approprier. A cet effet, l'enseignante B souhaite s'aider de la checklist fournie par le chercheur-intervenant, qui permet de s'assurer, en amont de la mise en œuvre avec les élèves, que l'artéfact numérique a été correctement paramétré pour que la catachrèse « coévaluation » puisse y opérer.

4.2.2 Focus groupe

Pour rappel, en fin d'année scolaire, l'année où a lieu la phase expérimentale et le recueil des données (année N), un focus groupe est organisé en visioconférence avec les trois enseignantes partenaires de cette recherche collaborative. L'enregistrement audio de ce focus groupe est consultable via un lien pérenne en Annexes 7.

En effet, chacune d'entre elles a étroitement collaboré avec le chercheur pour coconstruire, puis mettre en œuvre avec leurs élèves respectifs, le protocole semi-expérimental d'instrumentalisation de Moodle Atelier pour la coévaluation. En fin d'année scolaire, ce focus groupe permet donc de réunir les quatre acteurs de la présente recherche collaborative, en l'occurrence les trois enseignantes partenaires et le chercheur-intervenant – pour échanger à posteriori sur le vécu de chacun, de façon interactive. A ce stade, la phase expérimentale et le recueil des données sont terminés, toutefois la recherche collaborative se poursuit à travers ce focus groupe, avec de surcroît un objectif concret associant les trois enseignantes partenaires à la valorisation de la recherche, en l'occurrence un article co-écrit par elles et le chercheur (Brulin *et al.*, 2023) dans la revue d'interface *Les Cahiers pédagogiques* (voir Annexes 10).

Le premier thème abordé est celui de l'anonymat : il en ressort que les enseignantes ne voient pas d'inconvénient à ce que leur identité soit révélée, *e.g.* dans la section « Remerciements » ou dans le cadre de la publication de l'article, nécessairement signé. Les chefs d'établissement ont également donné leur accord pour que leur établissement soit mentionné. Une des idées qui est soulignée est le caractère valorisant de cette recherche collaborative dans le parcours professionnel des collègues, qui sont toutes titulaires du CAFFA et/ou impliquées dans un groupe de travail en lien avec leur rectorat d'académie. Par ailleurs, il est réaffirmé que les tâtonnements expérimentaux du chercheur ou des collègues partenaires sont inhérents à ce type de démarche et assumés en tant que tel – voire qu'ils peuvent être source d'inspiration pour la communauté éducative et scientifique. Comme précisé *supra*, l'anonymat des élèves est quant à lui intégralement protégé, conformément au formulaire de consentement libre et éclairé signé par leurs parents, et au règlement général sur la protection des données (RGPD) ce principe est réaffirmé en guise de préambule du focus groupe.

Nous rappelons ensuite la posture épistémologique qui caractérise la présente recherche, à savoir son inscription dans le paradigme de la recherche collaborative en sciences de l'éducation et de la formation, en précisant que si les collègues le souhaitent, outre le projet actuel de publication dans la revue d'interface *Les Cahiers Pédagogiques*, elles peuvent collaborer à d'autres projets de publication, ciblant des revues scientifiques cette fois.

Nous suggérons alors de faire un tour de table, pour que chacune des collègues partenaires puisse faire un retour sur expérience et que tous puissent en profiter pour réagir, en interaction. En parallèle, il est proposé d'adopter un cadrage thématique commun pour ces partages, issu en l'occurrence de l'approche instrumentale en environnement informatisé pour l'apprentissage humain. Ce n'est pas la première fois que nous rendons cette partie de notre cadre théorique explicite pour les enseignantes partenaires : précisément, le fait de partager ces construits théoriques avec elles en toute transparence est caractéristique de la posture épistémologique collaborative de cette recherche. Il s'agit en effet de les associer en toute conscience à la co-construction des savoirs, dans le cadre d'une recherche dont elles sont pleinement actrices – que ce soit lors de la phase exploratoire, lors de la phase expérimentale ou, comme indiqué précédemment, lors de la phase de valorisation.

L'enseignante A, pour commencer, souligne que l'axe technique/numérique n'est pas le point d'entrée qui l'intéresse le plus au départ ; en revanche, elle a été d'emblée particulièrement intéressée par l'artéfact pédagogique « coévaluation », en particulier parce qu'elle s'intéressait déjà à l'évaluation ; notamment comme moyen de mieux prendre en compte les difficultés des élèves porteurs de troubles des apprentissages, en s'écartant d'une approche strictement sommative, centrée sur la note et parfois stigmatisante, pour privilégier une approche par compétences (APC). La proposition de la présente recherche lui paraissait d'autant plus intéressante qu'il s'agissait de travailler cette APC de manière coopérative, d'une manière qui soit susceptible d'impliquer davantage les élèves dans une évaluation porteuse de sens – *e.g.* à travers la co-construction des grilles d'évaluation avec eux, et par l'entremise de l'artéfact pédagogique « coévaluation ». Sur le plan didactique, elle explique que les situations d'enseignement/apprentissage qu'elle a mises en œuvre dans le cadre de cette recherche ciblaient des compétences relatives à la production d'écrit et à la compréhension écrite, ainsi que les métacompétences en lien avec les critères d'évaluation afférents –notamment en vue de l'épreuve écrite de français du diplôme national du brevet (DNB).

Pour l'enseignante B en revanche, c'est le volet numérique du projet qui a emporté son adhésion en premier lieu, dans la mesure où depuis déjà environ cinq ans, elle s'intéresse de près aux innovations numériques susceptibles d'accompagner les situations d'enseignement/apprentissage qu'elle met en œuvre avec ses élèves. Le projet de recherche a été déployé auprès d'une classe de 3^e, cependant elle enseigne aussi auprès d'une classe comportant de nombreux élèves allophones et/ou non scolarisés antérieurement, avec une approche FLE. Dans son collège situé en réseau d'éducation prioritaire (REP), le numérique est donc vu comme une ressource supplémentaire, potentiellement engageante, pour favoriser

l'accès aux textes, aux mots, à l'écriture, et pour des élèves souvent en (grande) difficulté, afin d'essayer de réduire les inégalités socioscolaires.

L'enseignante B se dit également animée par le souci d'une évaluation plus juste et plus explicite. A ce titre, l'accompagnement proposé par le chercheur pour mettre en œuvre conjointement l'artéfact pédagogique « coévaluation », en instrumentalisant l'artéfact numérique « Moodle Atelier » lui a semblé intéressant, à fortiori pour bénéficier d'une fonctionnalité difficile – voire impossible – à mettre en œuvre sans le truchement du numérique, à savoir l'anonymisation en double aveugle. Il lui semble que le dispositif proposé permettait ainsi de pratiquer les pédagogies coopératives tout en contournant les difficultés liées aux conflits socioaffectifs entre élèves qui viennent d'ordinaire parasiter le travail en groupe dans sa classe. Lors de la première itération du protocole expérimental, ses élèves ont travaillé sur la production puis la coévaluation d'un écrit long. Ce choix correspond à ce que l'enseignante avait prévu de faire avant d'intégrer le projet ; il était donc lié à la nature écologique de cette recherche – en lien à la fois avec ses dimensions collaborative et semi-expérimentale. Un conflit instrumental a toutefois été généré pour un nombre significatif d'élèves : d'une part les critères d'évaluation (onze critères) – c'est-à-dire onze microcompétences travaillées (artéfact didactique) – étaient trop nombreux ; d'autre part les textes s'avéraient trop longs pour que les élèves disposent du temps suffisant pour coévaluer dans de bonnes conditions (artéfact pédagogique).

Pour la seconde itération, s'inspirant d'une activité mise en œuvre par l'enseignante A avec ses élèves de 3e, l'enseignante B a ainsi opté pour un écrit court, en réponse à une question de compréhension de type DNB. Le nombre de critères a par ailleurs été abaissé à six. Au dire de l'enseignante, cela s'est avéré plus fructueux, notamment en termes d'engagement des élèves dans les apprentissages – mieux mobilisés par le format court –, de temps consacré au projet – moins chronophage, plus resserré, donc plus cohérent – et de développement des compétences et métacompétences afférentes, grâce au nombre restreint de critères dans le formulaire de coévaluation.

L'enseignante C pratique pour sa part les pédagogies coopératives depuis plusieurs années au sein d'une « classe coop' ». Pour favoriser une dynamique d'engagement dans sa classe, elle se place volontiers dans une posture « d'enseignante-apprenante » : il s'agit d'entraîner les élèves à sa suite dans des apprentissages vécus comme une aventure collective, et dont elle joue le jeu au même titre qu'eux. A titre d'exemple, comme nous l'avons vu *supra*, elle participe – de façon anonyme comme ses élèves – à l'ensemble des ateliers d'écriture qu'elle a mis en œuvre dans le cadre du protocole expérimental, soumettant à chaque fois,

comme eux, une production écrite de sa composition, délibérément « imparfaite ». Ceci s'accomplit de manière transparente avec ses élèves, pour qui cela constitue en quelque sorte un jeu de piste : ceux qui doivent coévaluer la production de l'enseignante sauront-ils la déceler ? Elle confirme notamment que le dispositif expérimental a contribué à approfondir une pratique déjà en place avec ses élèves, à savoir la co-construction des grilles d'évaluation critériées. Elle souligne par ailleurs combien le partage du geste évaluatif avec les élèves, qui est au cœur même du dispositif, lui a permis de questionner ses pratiques hétéro-évaluatives à la lumière des pratiques coévaluatives des élèves, faisant évoluer ses schèmes d'enseignement en faveur d'une approche plus explicite, notamment en affinant son approche docimologique.

L'enseignante C revient en outre sur le rôle important que le dispositif a joué pour elle sur le plan didactique, en particulier pour un renouvellement de son approche de l'enseignement/apprentissage de la langue. En effet, « face aux copies des camarades, les élèves se sont mis à se poser des questions de grammaire, pour essayer de voir si le camarade était en conformité avec la consigne ; et, de fait, ils se sont posé des questions bien au-delà de ce que j'attendais ». Elle cite un exemple : lorsqu'il s'agissait d'introduire une comparaison dans une production écrite, les élèves se sont interrogés quant à d'autres figures d'analogie, « et ils se sont posé les bonnes questions ». Elle précise que ces questions émanaient le plus souvent d'élèves qui, bien qu'en difficulté, avaient à cœur d'évaluer avec justesse.

Partant de ces questions d'élèves, qui constituaient d'authentiques « situations-problèmes » au sens de (Barth, 2015), et adoptant la posture de l'enseignante-apprenante, l'enseignante C a pu accompagner les élèves dans ces apprentissages dont ils étaient eux-mêmes demandeurs. *In fine*, ces consolidations et approfondissements spontanés, issus de l'artéfact didactique d'une itération donnée, se sont naturellement retrouvés dans la grille de coévaluation pour l'itération suivante, conférant aux apprentissages une progressivité quasi organique.

Le témoignage de l'enseignante C concernant ses élèves de 5^e fait réagir l'enseignante B, dans la mesure où, pour leur part, ses élèves de 3^e n'ont pas semblé prendre conscience des enjeux métacognitifs liés à l'activité : était-ce en raison de leur âge ? L'enseignante A, en réponse, suggère que ce phénomène peut être davantage lié au profil de la classe (taux élevé d'élèves en grande difficulté), dans la mesure où elle a rencontré des difficultés similaires avec une classe de 4^e ayant le même profil que celui de l'enseignante B. La piste du climat de classe est également évoquée – piste qui semble corroborée par le fait que la classe de l'enseignante C est une « classe coop' », dotée d'un climat jugé très bon. Pour favoriser un climat de classe coopératif, propice à la coévaluation et au développement des métacompétences, les

enseignantes B et C s'accordent à dire que le dispositif pédagogique numérique de coévaluation gagnerait à être mis en œuvre de façon régulière.

L'enseignante C évoque également une gestion de classe facilitée grâce au dispositif, notamment du fait de l'anonymat. Elle signale par ailleurs que la mise en œuvre en salle informatique permet à chacun de fonctionner à son rythme, ce qui valorise le caractère mi-synchrone/mi-asynchrone et mi-présentiel/mi-distanciel du dispositif. En effet, dans l'espace de la salle informatique et pendant les 55 minutes que dure la séance, l'hybridité du dispositif favorise une certaine flexibilité pour les situations d'enseignement/apprentissage, propice à la différenciation pédagogique.

Les échanges portent ensuite sur la problématique de la note, « totem » qui tend à occulter les compétences évaluées – y compris, potentiellement, avec un dispositif doté d'une approche par compétences, comme celui expérimenté ici. Au dire de l'enseignante B, un conflit instrumental peut se faire jour ici, si aux yeux de l'élève l'artéfact pédagogique de coévaluation par compétences se réduit *in fine* à deux notes – c'est-à-dire la note de scripteur et la note de coévaluateur. En effet, dans ce cas l'accès à l'artéfact didactique peut se trouver restreint, c'est-à-dire que les compétences cibles et les métacompétences afférentes peuvent ne pas être suffisamment conscientisées par l'élève, pour des apprentissages pérennes. Comme stratégie de régulation pour « désacraliser » la fameuse note sur 20 et réhausser du même coup l'importance des compétences, autrement dit pour pallier cette potentielle combinatoire conflictuelle, l'enseignante A témoigne de la solution de compromis qu'elle a trouvée avec Moodle Atelier pour ses élèves : accommoder les schèmes d'utilisation de l'artéfact numérique en le paramétrant de telle sorte que la note soit exprimée sous forme de pourcentage.

4.2.3 Discussion (HO 4b)

En somme, les résultats obtenus permettent d'isoler trois grandes catégories de phénomènes en lien avec la genèse instrumentale en EIAH. Pour commencer, on constate des conflits instrumentaux affectant les situations d'enseignement. Ceux-ci se rapportent aux artéfacts didactique et technique du dispositif expérimental, ainsi qu'à une combinatoire pédagogique et technique. Le Tableau 20 résume les résultats obtenus.

Conflits instrumentaux observés	Artéfacts affectés
Accroissement des inégalités socioscolaires	Artéfact didactique
Choix des critères d'évaluation et prérequis didactiques	Artéfact didactique (renoncement partiel ou total)
Contresens didactique	Artéfact didactique (renoncement partiel)
Faible taux de participation à l'activité	Artéfact pédagogique
	Artéfact technique
La catachrèse « coévaluation »	Artéfact technique
Problèmes de connexion à l'artéfact numérique	Artéfact technique (retard d'apprentissage)
Manque de contrôle sur l'enseignement (non-accès au « Rapport d'évaluation »)	Artéfact technique

Tableau 20 – Principaux conflits instrumentaux observés et artéfacts afférents, du point de vue des enseignantes partenaires




Artéfact didactique	
Artéfact pédagogique	
Artéfact technique	

Tableau 21 – Légende

On peut récapituler ensuite les régulations instrumentales mises en œuvre. Ces dernières sont plus nombreuses que les conflits instrumentaux stricto sensu. Le Tableau 22 récapitule ces régulations instrumentales, en précisant pour chaque régulation observée les artéfacts conflictuels et les artéfacts régulateurs.

Régulations instrumentales observées	Artéfact(s) conflictuel(s)	Artéfact(s) régulateur(s)
Écrit long vs écrit court	Artéfact pédagogique	Artéfact didactique
Restriction du nombre de critères d'évaluation	Artéfact pédagogique	Artéfact didactique
Étayage métacognitif au moyen d'échelles de Likert et de descriptifs ad-hoc	Artéfact pédagogique	Artéfact didactique

Concertations spontanées entre élèves : débats docimologiques	Artéfact didactique	Artéfact pédagogique
Sentiment de perte de contrôle didactique et "rapport d'évaluation"	Artéfact didactique	Artéfact technique
« Totem » de la note et occultation des compétences cibles : pistes de régulation instrumentale	Artéfact pédagogique	Artéfact technique
La métacognition et la coopération comme enjeux d'une pratique régulière de la coévaluation	Artéfact pédagogique	Artéfact pédagogique
	Artéfact didactique	Artéfact technique
Artéfact technique, APC et réflexion docimologique	Artéfact pédagogique	Artéfact pédagogique
	Artéfact didactique	Artéfact technique
Ludification de la métacognition	Artéfact didactique	Artéfact pédagogique
		Artéfact technique
Augmentation de la motivation et de l'autorégulation des apprentissages des élèves en difficulté	Artéfact didactique	Artéfact pédagogique
		Artéfact technique
Instrumentation de l'étayage didactique par le dispositif pédagonumérique	Artéfact didactique	Artéfact pédagogique
		Artéfact technique
Meilleure compréhension des attendus didactiques	Artéfact didactique	Artéfact pédagogique
		Artéfact technique
Objectivation accrue des pratiques évaluatives	Artéfact pédagogique	Artéfact pédagogique
		Artéfact technique
Augmentation de la motivation intrinsèque	Artéfact pédagogique	Artéfact pédagogique
		Artéfact technique
Outiller le numérique pour une évaluation plus juste et plus explicite	Artéfact pédagogique	Artéfact pédagogique
		Artéfact technique
Hybridité du dispositif et différenciation pédagogique	Artéfact pédagogique	Artéfact pédagogique
		Artéfact technique

Tableau 22 – Régulations instrumentales observées, avec artéfacts conflictuels et artéfacts régulateurs afférents

Enfin, on constate que le dispositif expérimental de coévaluation a occasionné une évolution des schèmes d'utilisation des enseignantes partenaires, résumée dans le Tableau 23, en précisant à chaque fois l'artéfact concerné.

Evolution des schèmes d'utilisation des enseignants	Artéfacts concernés
Choisir et formuler les critères d'évaluation pour d'authentiques situations-problèmes	Artéfact didactique
Approche renouvelée de l'articulation entre étude de la langue et production d'écrit	Artéfact didactique
Objectivation accrue des pratiques évaluatives	Artéfact didactique
Horizontalité accrue de la posture enseignante en « classe coop' »	Artéfact pédagogique
Pistes de régulation instrumentale envisagées en présentiel synchrone ou distanciel asynchrone	Artéfact technique
Moodle Atelier : point d'entrée vers Moodle en général	Artéfact technique

Tableau 23 – Evolution des schèmes d'utilisation et artéfacts afférents

A titre d'exemple, la perte de contrôle ressentie par l'enseignante A était jugée par elle comme « ce qui est le plus problématique ici ». Elle a pu remédier à cela ensuite, mais cela a occasionné une perte de temps. On se trouve donc en présence d'un conflit instrumental se rapportant à l'artéfact technique et occasionnant un retard d'enseignement. Or il s'avère que l'artéfact numérique Moodle Atelier génère automatiquement en fin d'activité un fichier nommé « Rapport d'évaluation », qui récapitule, à destination de l'enseignant comme des apprenants, l'ensemble des éléments travaillés : production écrite, coévaluations des pairs et écrits commentatifs afférents, hétéro-évaluation de l'enseignant et écrit commentatif afférent, note de scripteur et note de coévaluateur. Toutefois, à ce stade précoce de la mise en œuvre de la présente étude, et du fait des contraintes liées à la situation écologique en général et à l'épidémie de Covid-19 en particulier, ces éléments n'avaient pas encore été portés à la connaissance de l'enseignante partenaire par le chercheur. En expliquant cet aspect du fonctionnement de Moodle Atelier à l'enseignante partenaire, le problème a pu être rapidement solutionné. Toutefois, la défaillance ici constatée en termes d'instrumentalisation de l'artéfact numérique (conflit instrumental d'origine technique), relativement à des schèmes d'utilisation ne nécessitant pourtant pas le recours à une instrumentalisation spécifique ou à une catachrèse – en l'occurrence le rapport d'évaluation –, génère une combinatoire conflictuelle suscitant un retard d'enseignement/apprentissage d'ordre pédagogique : temporairement, la coévaluation ne peut pas fonctionner pleinement, ni pour l'enseignante, ni pour les élèves. De surcroît, l'artéfact didactique s'en trouve affecté, puisqu'enseignant et apprenants sont temporairement privés de rétroactions pourtant nécessaires à la progression des apprentissages. En conséquence, ce conflit instrumental a dû faire l'objet d'une régulation, à savoir une vigilance

accrue lors de toutes les itérations ultérieures du protocole expérimental avec tous les échantillons.

On constate par ailleurs qu'au dire de l'enseignante C la coévaluation selon l'approche par compétences (APC), instrumentée par l'artéfact numérique pour la mise en œuvre de la coévaluation (artéfact pédagogique), a un effet rétroactif sur l'artéfact didactique. Relativement à la genèse instrumentale du dispositif, le triple enchâssement entre les artéfacts didactique, pédagogique et technique semble manifeste dans la mesure où l'utilisation de l'artéfact numérique – et notamment de son interface graphique – a instrumenté la coévaluation (artéfact pédagogique) de telle sorte que l'APC puisse y catalyser la réflexion docimologique – c'est-à-dire métacognitive – des élèves participants, relativement aux compétences cibles (artéfact didactique). En outre, l'initiation d'un processus de réflexion docimologique en classe semble pouvoir constituer un facteur de régulation pour plusieurs conflits instrumentaux identifiés *supra*, en l'occurrence les « problèmes de discernement docimologique », la « vision conflictuelle entre inter- et hétéro-évaluation » et les « inhibitions socioaffectives » (*e.g.* la peur de noter injustement un camarade de classe).

De même, l'artéfact pédagogique « Coévaluation » instrumente l'artéfact didactique « Etude de la langue », dans la mesure où il incite l'enseignante B à modifier ses schèmes d'utilisation en réduisant le nombre de critères évalués, par rapport à ce qu'elle pratique d'habitude. Par ailleurs, en dehors de toute considération liée à la présente recherche collaborative, la mise en œuvre d'une séance de coévaluation par l'entremise de l'artéfact Moodle Atelier semble générer un conflit instrumental d'origine technique/numérique, dans la mesure où l'appropriation de cet artéfact est un processus complexe avec, concrètement, de nombreuses opérations de paramétrage nécessaires à son instrumentalisation – autrement dit pour effectuer la catachrèse détaillée *supra*. Ceci étant, le fait que la recherche collaborative contribue à l'opérationnalisation de la différenciation pédagogique (DP) dans la classe en nourrissant la démarche réflexive des enseignantes partenaires à l'égard des pratiques évaluatives – *e.g.* choix des compétences cibles, rationalisation et transparence accrues, nuances docimologiques – semble confirmer les préconisations de (Forget, 2017b) concernant l'adéquation entre le paradigme collaboratif de la recherche et les défis posés par la DP, à la fois en termes d'opérationnalisation dans la recherche et de mise en œuvre dans les pratiques de classe.

On constate en outre que, lors de la première itération du protocole expérimental avec la classe B, une liste de dix critères – en l'occurrence dix microcompétences linguistiques – constitue la grille de coévaluation. Un conflit instrumental d'origine didactique peut ici affecter

la coévaluation numérique (artéfact pédagogique), dans la mesure où le dispositif pourrait bénéficier surtout aux élèves déjà dotés de compétences linguistiques solides, tandis que les élèves les moins compétents sur ce plan ne pourraient probablement pas repérer les erreurs linguistiques de leurs pairs ou, à tout le moins, disposer en la matière d'un sentiment d'efficacité personnelle (Bandura, 2007) et/ou d'une motivation suffisante (Cosnefroy & Fenouillet, 2009) pour adopter une posture de coévaluateur. Si les élèves déjà en réussite continuent à progresser grâce au dispositif, les écarts entre élèves pourraient s'accroître. Cependant nous avons vu avec Forget (2017b) que les progrès des élèves les plus en réussite deviennent marginaux au-delà d'un certain seuil de maîtrise. D'autre part, en appui sur la dimension coopérative du dispositif, notamment par le truchement des écrits commentatifs, les élèves les plus en difficulté peuvent théoriquement progresser, même s'ils sont dans la situation de maîtrise insuffisante précitée.

De même, des débats docimologiques – à l'instar de celui qui s'est déroulé spontanément au CDI du collège C à T1/T2 – semblent à même de potentialiser, dans le dispositif expérimental de coévaluation, une dimension coopérative latente. Non seulement le dispositif est coopératif par nature, puisqu'il implique une coopération entre l'enseignant et ses élèves, qui « partagent l'évaluation » de manière synchrone et présentielle, quoique par ordinateur interposé ; mais de surcroît on s'aperçoit que le dispositif peut se décliner sous forme groupale, *e.g.* avec des îlots de coévaluateurs qui coopèrent pour une évaluation plus juste. Cette coopération peut avoir lieu en présentiel – comme dans la scène qui s'est déroulée au CDI – ou en distanciel, *e.g.* par messagerie interposée, par téléphone, via les réseaux sociaux, dans le cas d'un aménagement en tout distanciel, tel que celui envisagé par l'enseignante B.

Dans plusieurs témoignages d'enseignantes partenaires, une dimension motivationnelle apparaît par ailleurs. Elle a trait à une forme de ludification de la coévaluation, à savoir le fait pour les élèves de « prendre la place du prof », singulièrement eu égard à la rédaction d'écrits commentatifs à destination des pairs. Ce constat est largement corroboré par les résultats issus des données recueillies auprès des élèves participants. De surcroît, des phénomènes relevant de l'autorégulation des apprentissages ont pu être observés, en lien avec les notions connexes de métacognition et de motivation. En effet, au dire de l'enseignante C notamment, le passage d'une contrainte vécue comme extrinsèque, à une appropriation intrinsèque de celle-ci, est vecteur de motivation pour les élèves, et s'appuie sur une prise de recul de ces derniers vis-à-vis de leurs apprentissages – c'est-à-dire un accroissement de leurs capacités métacognitives : la contrainte est acceptée et intériorisée par l'élève car il comprend qu'elle est source de créativité. Ce processus est susceptible d'avoir un effet positif sur les compétences des élèves

en termes de maîtrise de la langue et de production d'écrit (artéfact didactique). Il prend notamment racine, d'après l'enseignante C, dans la pratique de la coévaluation (artéfact pédagogique), que l'instrumentalisation de Moodle Atelier (artéfact numérique) a facilitée – voire rendue possible, singulièrement si l'on considère certaines fonctionnalités inaccessibles sans le recours à un tel artéfact, *e.g.* anonymisation en double aveugle, note de coévaluateur.

Dans Moodle Atelier, deux fonctionnalités avancées permettraient de pousser encore plus loin les effets positifs de ces synergies régulatrices sur la genèse instrumentale du dispositif de coévaluation numérique. La première de ces fonctionnalités, détaillée *supra*, a été exploitée pour l'échantillon A à T4 : la possibilité de fournir des « Exemples de travaux ». L'autre n'a été exploitée avec aucun échantillon ; il s'agit des « Travaux publiés », c'est-à-dire la possibilité de publier une sélection de travaux d'élèves dans l'atelier Moodle, anonymement ou non. En effet, ces fonctionnalités permettent de mettre en exergue certains faits de langue à même des productions écrites d'élèves, c'est-à-dire *in situ*, à partir d'écrits auxquels les élèves peuvent s'identifier d'autant mieux qu'ils ont été produits par des pairs. En particulier, les « Travaux publiés » peuvent accompagner la mise en œuvre d'une remédiation, la consolidation d'une compétence, voire la scénarisation d'un approfondissement notionnel. De plus, de tels scénarios peuvent être mis en œuvre « en temps réel », notamment en se fondant sur des écrits que les élèves viennent de produire, ou en puisant dans les archives anonymisées des années précédentes – que Moodle permet de sauvegarder. Le potentiel de l'artéfact numérique s'avère donc significatif pour la mise en œuvre de la différenciation pédagogique : il peut outiller efficacement l'artéfact pédagogique « coévaluation », en vue d'une approche renouvelée, différenciée voire personnalisée, de l'artéfact didactique – mais dans la perspective coopérative (Connac & Cieutat, 2021 ; Reynaud, 2022) d'un collectif apprenant (Lescouarch, 2021). Cette notion de collectif apprenant, qui déborde le groupe classe pour inclure l'équipe enseignante, peut également impliquer la collaboration entre praticiens et chercheurs, comme c'est le cas dans la présente recherche. D'autre part, pour favoriser un climat de classe propice à la coopération, les enseignantes B et C ont clairement exprimé leur volonté de pratiquer fréquemment la coévaluation numérique dans leurs classes. En tant que praticiennes, leur intuition professionnelle corrobore donc les conclusions de Quesada *et al.* (2019) relativement à la nécessité de mettre régulièrement en œuvre ce type de pratique coévaluative dans le quotidien de la classe, afin d'en optimiser les apports.

En somme, nos résultats permettent de valider l'hypothèse HO 4b, c'est-à-dire « La coévaluation numérique génère des conflits instrumentaux affectant les situations d'enseignement, mais ces derniers peuvent être régulés ».

4.3 Discussion (HO 4)

Les données textuelles en provenance des élèves participants et des enseignantes partenaires permettent de mettre en évidence un nombre significatif de conflits instrumentaux (CI) affectant les situations d'apprentissage et/ou d'enseignement. Singulièrement, l'analyse a montré qu'il s'agit toujours de combinaisons conflictuelles dans la mesure où, au sein du triple enchâssement des artefacts didactique, pédagogique et technique en EIAH, les CI observés se rapportent toujours à plusieurs artefacts.

De surcroît, dans le prolongement des travaux de Marquet (2005b), Marquet & Coulibaly (2011) et Zeller & Marquet (2021b), en appui sur une typologie permettant de catégoriser ces conflits instrumentaux selon leur degré de conflictualité – autrement dit selon qu'ils entraînent un retard, un renoncement partiel ou un renoncement total –, on constate que les combinaisons conflictuelles observées peuvent avoir des conséquences significatives sur les situations d'enseignement et/ou sur les situations d'apprentissage.

Toutefois, dans la plupart des cas, ces combinaisons conflictuelles s'accompagnent de régulations instrumentales. Plus précisément, il s'agit de synergies régulatrices, impliquant là aussi des combinaisons de facteurs de régulation, liés aux artefacts didactique ou pédagogique – ou, dans le cas de l'artefact technique, des combinaisons de fonctionnalités régulatrices qui sont l'apanage du numérique. Or nos résultats montrent par ailleurs que les couples conflit/régulation sont plus nombreux que les conflits instrumentaux stricto sensu. Dans l'ensemble, cette dynamique de génération/résolution de conflits instrumentaux – autrement dit la dialectique entre combinaisons conflictuelles d'une part, et synergies régulatrices d'autre part – semble donc à même de préserver le potentiel de la coévaluation numérique en termes de mise en œuvre de la différenciation pédagogique en classe.

En somme, l'analyse des données issues des élèves participants et des enseignantes partenaires permet de valider l'hypothèse HO 4, c'est-à-dire « La coévaluation numérique génère des conflits instrumentaux, mais ces derniers peuvent être régulés ».

PARTIE V : DISCUSSION GENERALE

Dans une perspective transversale, cette partie a pour objectif d'effectuer un retour réflexif sur notre cadre théorique, en combinant les résultats liés à nos quatre hypothèses opérationnelles (Tableau 24).

	Hypothèses opérationnelles	Validation
HO 1	Amélioration des compétences en français	Oui
HO 2	Amélioration de l'autodirection des apprentissages	Oui
HO 3	Amélioration des apprentissages coopératifs	Oui
HO 4	Génération et régulation de conflits instrumentaux	Oui

Tableau 24 – Récapitulatif des quatre hypothèses opérationnelles et de leur validation

Nous reviendrons dans un premier temps sur les apports du dispositif expérimental de coévaluation en termes d'apprentissages en français langue première (chapitre 1). Nous nous interrogerons quant à la capacité du dispositif à promouvoir l'autodirection des apprentissages (chapitre 2), avant de questionner les apports de ce dernier sur le plan des apprentissages coopératifs (chapitre 3). Nous effectuerons ensuite un retour réflexif sur l'ensemble de nos résultats à l'aune de l'approche instrumentale (chapitre 4). Dans les deux chapitres suivants, nous nous demanderons dans quelle mesure l'opérationnalisation de la différenciation pédagogique est rendue possible : sur le terrain, par la coévaluation numérique (chapitre 5) et, dans la recherche, au moyen du paradigme collaboratif (chapitre 6). Pour finir, nous examinerons les limites de la présente étude (chapitre 7).

1 Coévaluation numérique, maîtrise de la langue et production d'écrit

Nos résultats ont permis de valider l'hypothèse HO 1, c'est-à-dire « La coévaluation numérique améliore les compétences des élèves en français langue première », en termes de maîtrise de la langue et de production d'écrit. En convoquant à nouveau le cadre théorique de référence, ce chapitre a pour objectif de développer la discussion en croisant l'ensemble de nos résultats – autrement dit en mobilisant les résultats issus de nos quatre hypothèses opérationnelles.

1.1 Contresens pédagogique et empêchement didactique

Nous avons pu identifier un certain nombre de conflits instrumentaux qui, dans la genèse instrumentale du dispositif, gênent l'accès à l'artéfact didactique. Notamment lorsque certains élèves ne semblent pas disposer du prérequis pédagogique nécessaire pour comprendre la signification ou l'utilité de la coévaluation. Les élèves concernés par ce contresens pédagogique à l'égard de la coévaluation semblent avoir « manqué la cible » sur le plan pédagogique avec, à la clé, d'éventuels renoncements partiels – voire totaux – à la tâche d'apprentissage, gênant ou empêchant l'accès à l'artéfact didactique. Les verbatim d'élèves présentés dans le Tableau 25 montrent cependant une réalité plus nuancée. Pour cet échantillon restreint, nous ne pouvons pas établir de corrélation au moyen d'une analyse inférentielle (corrélation de Spearman, $p > .05$) ; on constate simplement que les profils sont très hétérogènes et qu'aucun schéma d'ensemble ne se dégage – y compris lorsque l'on considère un indicateur de tendance centrale, *e.g.* moyenne. Pour un échantillon un peu plus large, en l'occurrence le super-échantillon SE (N = 65), nous avons toutefois pu établir une corrélation positive à T4 ; en d'autres termes, lorsque le score de coévaluateur est élevé, le score de scripteur tend à l'être aussi. Réciproquement, lorsque le score de coévaluateur est bas, le score de scripteur tend à l'être aussi. Ainsi les résultats de nos analyses quantitatives sur données numériques corroborent les résultats qualitatifs issus de nos données textuelles, et confirment qu'un contresens pédagogique est susceptible d'entraîner un empêchement didactique.

Elève	La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim)	Moyenne coévaluateur (%)	Moyenne scripteur (%)
A	Facile et difficile à la fois	C'était plutôt long mais ça va c'était simple	61	42
B	Facile	C'est facile à faire	87	50
C	Facile	Simple	62	61
D	Facile	Parce qu'il faut juste évaluer un truc	62	44
E	Très facile	Ce n'était pas très compliqué	31	33
F	Très facile	Parce que c'est facile.	76	89
G	Très facile	Personnellement je ne suis pas très indulgent, donc c'était assez facile pour moi	45	72
H	Très facile	Oui c'était facile mais je trouve ça très nul : cela va me servir à rien dans ma vie	65	52
I	Très facile	Parce que c'est un QCM sans mauvaises réponses.	77	92
		Moyennes	63	59

Tableau 25 – Contresens pédagogique et empêchement didactique : verbatim d'élèves et scores moyens de scripteurs et de coévaluateurs (T2-T4, en %)

1.2 Prérequis didactiques et conflit instrumental pédagogique

A l'instar du prérequis pédagogique, la notion de prérequis didactique semble importante dans la mesure où les élèves qui manifestent déjà une bonne maîtrise des compétences cibles sont susceptibles de développer une maîtrise plus grande encore, tandis que ceux qui ne possèdent pas les prérequis didactiques nécessaires pour s'engager dans cette tâche

métacognitive, donc exigeante, qu'est la coévaluation, ne semblent pas pouvoir bénéficier pleinement de la situation d'apprentissage.

Or lorsqu'on s'intéresse aux scores de coévaluateurs des élèves ayant obtenu les scores de scripteurs les moins élevés (score inférieur ou égal à la valeur du 3^e quartile), on constate qu'en moyenne les scores de coévaluateurs sont deux fois plus élevés que les scores de scripteurs, à T2 comme à T4 (Tableau 26 et Tableau 27). Nous avons établi par ailleurs que les scores de scripteurs de l'ensemble des élèves participants ont globalement augmenté entre T2 et T4. Cela est également vrai dans le cas des élèves ayant obtenu une note de scripteur inférieure ou égale à la valeur du 3^e quartile, puisque ces scores de scripteurs connaissent une hausse légère mais significative (+4%). Sachant que nous avons pu établir par ailleurs une corrélation entre les scores de scripteurs et les scores de coévaluateurs pour l'ensemble des élèves à T4, on peut déduire de tout ce qui précède que, dans ce contexte de hausse des scores de scripteurs à T4, et vu la corrélation avec les scores de coévaluateurs, le dispositif pédagogique de coévaluation contribue à une meilleure maîtrise des compétences cibles, par le truchement de la métacognition. En d'autres termes, lorsqu'à T2 les élèves ayant obtenu un faible score de scripteur ont évalué leurs pairs, ils sont montés en métacompétences, ce qui a entraîné une montée en compétences à T4.

Code élève	Scores de scripteurs (T2, en %)	Scores de coévaluateurs (T2, en %)
SET23Q1	8	41
SET23Q2	17	50
SET23Q3	17	26
SET23Q4	22	70
SET23Q5	25	77
SET23Q6	27	48
SET23Q7	27	79
SET23Q8	27	60
SET23Q9	28	30
SET23Q10	28	43
SET23Q11	28	46
SET23Q12	30	71
SET23Q13	33	78
SET23Q14	33	70
SET23Q15	33	31
SET23Q16	33	76
SET23Q17	33	77
SET23Q18	33	52
Moyenne	27	57

Tableau 26 – Scores de scripteurs et de coévaluateurs des élèves du 3e quartile à T2 (par score de scripteur croissant, en %)

Code élève	Scores de scripteurs (T4, en %)	Scores de coévaluateurs (T4, en %)
SET43Q1	0	44
SET43Q2	9	33
SET43Q3	25	33
SET43Q4	26	89
SET43Q5	35	56
SET43Q6	40	33
SET43Q7	40	56
SET43Q8	40	67
SET43Q9	42	33
SET43Q10	42	67
SET43Q11	44	78
Moyenne	31	54

Tableau 27 – Scores de scripteurs et de coévaluateurs des élèves du 3e quartile à T4 (par score de scripteur croissant, en %)

Rappelons que cette tendance concernant les élèves ayant obtenus les scores de scripteurs les moins élevés se retrouve plus globalement parmi l'ensemble des élèves participants. En particulier, la corrélation positive que nous avons pu établir à T4 entre scores de scripteurs et scores de coévaluateurs confirme l'importance des prérequis didactiques, à la fois pour performer en tant que scripteur et pour coévaluer efficacement : lorsque le score de scripteur est bas – ce qui indique potentiellement une absence de prérequis didactiques –, le score de coévaluateur tend à l'être aussi, probablement en raison d'une difficulté à entrer dans la métacognition (métacompétences) à propos de compétences insuffisamment maîtrisées. Néanmoins, nous avons vu que ce conflit instrumental d'origine didactique, qui affecte l'artéfact pédagogique, peut également (ou alternativement) avoir pour origine des croyances motivationnelles faibles (Cosnefroy & Fenouillet, 2009), potentiellement en lien avec un sentiment d'efficacité personnelle peu élevé (Bandura, 2007).

1.3 Problèmes de discernement docimologique

Un nombre significatif d'élèves ont déclaré rencontrer des difficultés à évaluer la production écrite d'un pair, notamment en raison de problèmes de discernement docimologique, à savoir la confrontation à un dilemme lorsqu'il s'agit de déterminer quel est

le juste niveau de maîtrise, relativement à telle ou telle compétence coévaluée dans la production écrite d'un pair. Là encore, on peut supposer une maîtrise insuffisante des compétences cibles, qui fait écho à la potentielle absence de prérequis didactiques précédemment évoquée pour certains élèves. Ce conflit instrumental d'origine didactique est susceptible d'entraîner un retard, voire un renoncement partiel, à l'égard de l'artéfact pédagogique, mais aussi à l'égard de la composante métacognitive – c'est-à-dire les métacompétences cibles – de l'artéfact didactique. Les verbatim présentés dans le Tableau 28 rendent compte de ce phénomène, en précisant, pour l'auteur de chaque verbatim, le score moyen de scripteur et de coévaluateur pour l'ensemble de la période expérimentale. L'échantillon rapporté ici ne se prête pas à une analyse systématique, d'autant plus que les profils sont très variables et que les scores de scripteurs et de coévaluateurs ne semblent pas corrélés. Globalement, ces verbatim d'élève témoignent néanmoins concrètement de la prégnance de ces problèmes de discernement docimologique, attribuables à des difficultés d'ordre didactique.

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim)	Moyenne scripteur (%)	Moyenne coéval. (%)
Facile et difficile à la fois	Je trouve que c'est compliqué de faire un jugement sur un texte	42	84
Facile et difficile à la fois	C'était un peu dur de savoir comment évaluer.	67	73
Facile et difficile à la fois	Le travail à faire était facile à comprendre mais au début j'avais un peu de mal à donner mon avis sur le travail de mes camarades	97	88
Facile	C'était assez facile, mais assez compliqué à la fois, car on ne sait pas comment bien noter...	52	72
Facile et difficile à la fois	Parce que j'ai très rarement évalué d'autres personnes que moi, donc c'était difficile mais d'un autre côté, certaines consignes étaient faciles à évaluer	59	66
Facile et difficile à la fois	Je ne savais pas tout le temps quelle case cocher car il y avait des trucs bien et moins biens à la fois.	62	52
Très facile	Il n'y a pas de milieu entre « Fragile » et « Maitrisé », du coup on est obligé de trancher...	92	77
Facile et difficile à la fois	Car il faut à la fois être juste et non en même temps	21	48
Plutôt facile	L'hésitation entre deux cases, donner des bons conseils et que les autres les suivent	84	60
Plutôt facile	C'était bien mais difficile de savoir si notre évaluation est juste	59	51
Facile et difficile à la fois	Parce que c'est difficile de savoir mettre les bonnes notes pour évaluer justement un texte	94	72

Tableau 28 – Problèmes de discernement docimologique : verbatim d'élèves et scores afférents (en %)

1.4 Combinatoire conflictuelle en lien avec les interfaces de l'artéfact numérique

Nous avons vu que, dans le dispositif expérimental de coévaluation numérique, la qualité linguistique des productions d'écrit joue un rôle charnière pour l'articulation des artéfacts pédagogique, didactique et technique. En effet, le fait que les écrits soient produits sur ordinateur (sur traitement de texte), implique théoriquement que les élèves disposent de l'aide automatisée d'un correcticiel. Toutefois, comme détaillé *supra*, un tel logiciel peut faire l'objet d'un non-usage ou de mésusages, singulièrement pour des élèves en difficulté. De même, une interface comme le clavier peut être la source de conflits instrumentaux d'origine technique affectant en retour l'artéfact didactique, *e.g.* lorsqu'on ne sait pas où se trouve un caractère spécial ou un accent. Il en résulte une combinatoire conflictuelle. D'une part, ces mésusages ou non-usages sont susceptibles d'avoir un effet négatif sur la qualité linguistique des productions écrites : on se trouve en présence d'une combinatoire conflictuelle impliquant l'articulation entre artéfacts didactique et technique, et qui entraîne un renoncement partiel à la tâche d'apprentissage (artéfact didactique) pour l'élève scripteur – surtout pour les apprenants dont les compétences linguistiques sont les plus fragiles. D'autre part, ces mésusages ou non-usages de certains élèves scripteurs sont fondés à rendre la tâche singulièrement plus difficile pour le pair coévaluateur, qui se trouve *in fine* confronté à des écrits où les erreurs linguistiques surabondent, et qui sont donc très difficiles ou longs à déchiffrer/évaluer. Cela peut entraîner pour le pair coévaluateur un retard d'apprentissage – voire un renoncement partiel – vis-à-vis de la tâche de coévaluation (artéfact pédagogique), et par voie de conséquence, vis-à-vis de l'artéfact didactique également : en effet, la tâche métacognitive afférente aux compétences cibles en maîtrise de la langue peut elle aussi s'en trouver affecter. La Figure 72 illustre cette combinatoire conflictuelle.

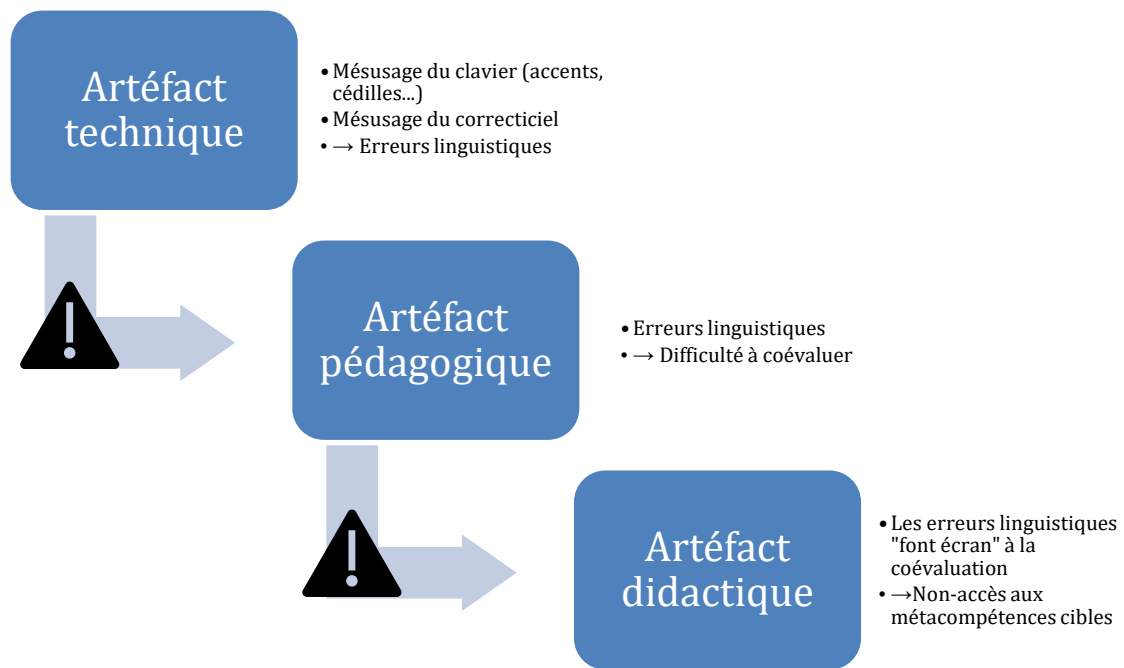


Figure 72 – Mésusages ou non-usages des interfaces associées à l'artéfact numérique : combinatoire conflictuelle

1.5 Synergie régulatrice favorisant la motivation intrinsèque, en lien avec l'artéfact didactique

Si la genèse instrumentale du dispositif expérimental de coévaluation numérique suscite des conflits instrumentaux, on constate par ailleurs que des processus de régulation instrumentale peuvent advenir de manière concomitante. Parfois, ces processus de régulation mobilisent une combinatoire d'artéfacts aboutissant à des synergies régulatrices. Le cas de Marie, élève de l'échantillon C, est intéressant à cet égard. Nous avons vu qu'elle a éprouvé de la difficulté pour intérioriser une consigne d'écriture – c'est-à-dire une micro-compétence cible –, en l'occurrence dans le cadre d'un travail sur les figures d'analogie en général, et sur la figure stylistique de la métaphore en particulier. L'absence de motivation intrinsèque face à une consigne d'écriture dont elle ne comprenait pas l'intérêt (contresens didactique partiel) a entraîné un conflit instrumental didactique. A la faveur du dispositif expérimental de coévaluation numérique, Marie semble cependant avoir bénéficié de l'emboîtement des artéfacts pédagogique et numérique, pour un accès facilité à l'artéfact didactique. En effet, l'analyse des données textuelles recueillies auprès d'elle nous a permis de constater qu'elle est passée, par l'entremise de la métacognition et de la coopération, d'une approche extrinsèque à une approche intrinsèque de la consigne d'écriture. En d'autres termes, en découvrant et en coévaluant les écrits de ses pairs, Marie s'est trouvée en contact avec des exemples concrets,

illustrant la consigne d'écriture, et auxquels elle pouvait s'identifier. On peut parler ici d'inspiration, mais aussi d'émulation entre pairs.

Cela est susceptible d'avoir entraîné pour elle un surcroît de motivation, et d'avoir renforcé chez elle la capacité à autoréguler ses apprentissages, à fortiori dans la mesure où l'enseignante C offrait à ses élèves l'opportunité de retravailler leurs textes pour prendre en compte les différents écrits commentatifs reçus, avant une nouvelle hétéro-évaluation de l'enseignante – sommative, cette fois. A priori, cette synergie régulatrice entre artéfacts pédagogique et numérique serait d'autant plus à même de bénéficier aux apprentissages qu'elle serait instrumentalisée d'emblée dans le dispositif pédagogique numérique de coévaluation. Comme indiqué *supra*, une catachrèse ad hoc permettrait de mettre en œuvre, à même l'artéfact numérique, une boucle de rétroaction formative anonymisée permettant de retravailler les écrits afin de les améliorer, avant de les soumettre à nouveau à l'épreuve – formative voire, *in fine*, sommative – de la coévaluation.

L'analyse quantitative des données numériques concernant Marie corrobore ces observations, à savoir que le dispositif pédagogique numérique de coévaluation peut contribuer à réguler des conflits instrumentaux liés à l'artéfact didactique. En effet, bien que son score à l'échelle SEP(srl)⁶² soit resté stable de T1 à T6 à un niveau moyen – c'est-à-dire avec un score individuel médian de 3 sur une échelle de Likert allant de 1 à 5 –, ses performances de scripteur ont été excellentes et en constante amélioration de T2 à T6 (Figure 73).

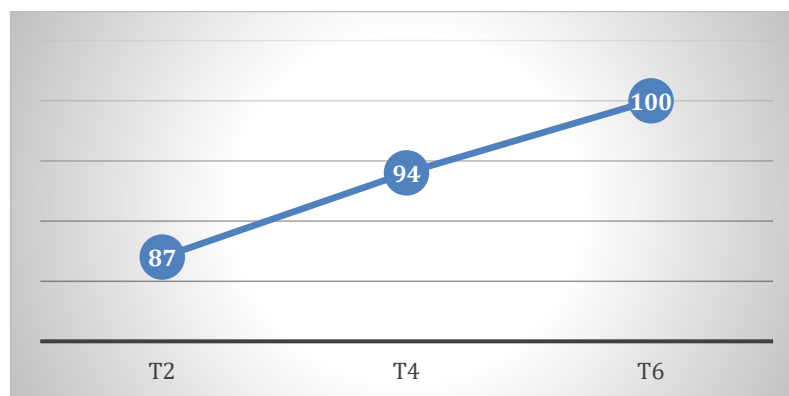


Figure 73 – Evolution du score de scripteur de Marie de T2 à T6 (en %)

De même, ses performances de coévaluatrice ont connu un accroissement significatif et constant de T2 à T6 (Figure 74) : partant d'un niveau moyen (55%), Marie a atteint un niveau élevé (84%).

⁶² SEP(srl) : Sous-échelle de l'échelle SEPE (sentiment d'efficacité personnelle des enfants/adolescents) mesurant l'auto-efficacité pour l'autorégulation des apprentissages.

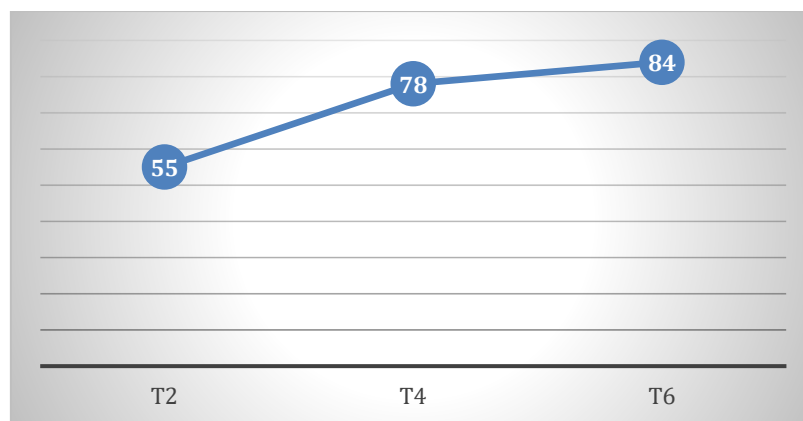


Figure 74 – Evolution du score de coévaluatrice de Marie de T2 à T6 (en %)

Or nous avons établi par ailleurs que, pour la période expérimentale commune à l'ensemble des échantillons (T1-T4), non seulement les scores de scripteurs, mais aussi les scores de coévaluateurs sont en hausse significative pour le super-échantillon (SE). Replacé dans ce contexte, le cas de Marie permet de supposer que l'amélioration globalement constatée eu égard aux scores de scripteurs et de coévaluateurs est liée aux apports de la coévaluation numérique (artéfacts pédagogique et numérique) en termes d'amélioration des compétences métacognitives et des apprentissages coopératifs, favorisant le développement d'une motivation intrinsèque.

1.6 Le dispositif pédagogique numérique : source à la fois de conflits et de régulations didactiques

Lors de la seconde réappropriation du dispositif par l'enseignante B en année N+1, nous avons observé un contresens didactique (conflit instrumental didactique), qui consistait pour certains élèves en un renoncement partiel à la tâche d'apprentissage : en effet, en guise de production écrite, ces derniers effectuaient des copiés-collés issus du web. A cet égard, nous avons vu que Noé⁶³ – un cas flagrant de plagiat intégral – obtenait un score de scripteur très élevé lorsqu'on considérait uniquement les 5 inter-évaluations issues de ses pairs (85%) : en effet, ces derniers n'avaient pas pu prendre en compte le facteur plagiat dans leur évaluation, dans la mesure où il s'agissait d'un imprévu didactique.

⁶³ Comme indiqué *supra*, les prénoms d'élèves ont été changés pour préserver leur anonymat, conformément au formulaire de consentement libre et éclairé signé par leurs parents et conformément au règlement général de protection des données (RGPD).

Cependant, dans l'hétéro-évaluation de l'enseignante, ce facteur a été pris en compte : un score de 55% lui a été attribué – qui visait selon l'enseignante B un compromis entre pénalisation du plagiat et valorisation de l'effort, certes limité, fourni par un élève habituellement peu ou pas engagé. Surtout, cette dernière a eu soin de configurer Moodle Atelier conformément au protocole co-construit en année N – de telle sorte que sa note ait le coefficient adéquat pour que la catachrèse « coévaluation » puisse opérer –, ce qui a permis de pondérer la surnotation des pairs et d'aboutir pour Noé à une note finale de scripteur ajustée, issue de l'agrégation pondérée entre les cinq inter-évaluations des pairs et l'hétéro-évaluation de l'enseignante, en l'occurrence 60%.

En outre, le contresens didactique qu'on peut supposer pour Noé – c'est-à-dire le conflit instrumental ayant entraîné un renoncement partiel à l'activité d'apprentissage et empêché, au moins en partie, son accès à l'artéfact didactique – a été rendu manifeste par ses résultats inférieurs à la moyenne quant à sa propre auto-évaluation et aux coévaluations qu'il a proposées pour les travaux de ses pairs. En effet, ses scores d'auto-évaluateur et de coévaluateur étaient tous deux de 45%.

Quant aux pairs qui ont largement surévalué sa production écrite car ils ignoraient qu'il s'agissait d'un plagiat, ils ont été pénalisés dans leur score d'évaluateur (45% en moyenne). En d'autres termes, leurs scores d'évaluateurs ne sont pas nécessairement révélateurs de leurs métacompétences réelles, relativement aux critères de réussite pour la réponse écrite à une question de compréhension type DNB (l'artéfact didactique ici). La « dissonance métacognitive » qui peut en résulter constitue un autre conflit instrumental, susceptible d'avoir entraîné là encore un renoncement partiel à l'activité d'apprentissage et empêché, au moins en partie, l'accès de ces coévaluateurs d'élèves plagiaires à l'artéfact didactique. Signalons toutefois que ce dernier conflit instrumental a pu être partiellement régulé par l'enseignante B, en ne comptabilisant pas les notes d'évaluateurs en question.

Il paraît toutefois intéressant de constater que, si l'instrumentalisation de l'artéfact pédagogique « coévaluation » au moyen de l'artéfact numérique « Moodle Atelier » a ici suscité des conflits instrumentaux – empêchant partiellement l'accès des élèves à l'artéfact didactique –, cette même instrumentalisation, et singulièrement la catachrèse afférente, ont permis de réguler ces conflits. Selon les termes employés par (Zeller & Marquet, 2021, p. 9), on constate donc effectivement que l'instrumentalisation de l'artéfact numérique « [génère] des conflits instrumentaux tout en permettant paradoxalement de les gérer, dans un processus d'ajustement ».

1.7 Lecture et coévaluation individualisée outillée par le numérique

Le dispositif de coévaluation numérique expérimenté dans le cadre de la présente recherche implique, de fait, une activité de lecture de la part du pair coévaluateur. Nous avons vu par ailleurs que ce dispositif met en œuvre une relation d'aide réversible, dans le cadre de laquelle chaque élève est simultanément scripteur inter- et hétéroévalué, et coévaluateur de plusieurs écrits de pairs. Ainsi, tous les élèves sont amenés à mobiliser leur métacognition pour jouer ce rôle de lecteur attentif des écrits de leurs pairs.

En termes de motivation extrinsèque, cette activité de lecture est, à priori, d'autant plus attentive qu'elle donne lieu à un score de coévaluateur⁶⁴. Certes, nous avons vu que pour des élèves ayant un profil de quasi-décrocheurs scolaires – *e.g.* certains élèves de l'échantillon B, dont le collège est situé en réseau d'éducation prioritaire (REP) –, ce facteur de motivation extrinsèque qu'est le score de coévaluateur peut ne pas suffire à les engager dans la tâche coévaluative. Toutefois, nous avons eu en « classe coop' » (échantillon C) un contre-exemple, avec un élève en grande difficulté socioscolaire qui s'est néanmoins fortement investi dans l'activité proposée.

En termes de motivation intrinsèque, cette activité de lecture est d'autant plus porteuse qu'elle offre aux élèves l'opportunité de découvrir les écrits de leurs pairs, ce dont ils ont souvent déclaré se réjouir. Ils soulignent également que la lecture des écrits de leurs pairs est pour eux une source potentielle d'inspiration ou d'émulation, comme nous l'avons vu à travers de nombreuses données textuelles recueillies auprès des élèves participants ou, *supra*, à travers l'exemple de Marie et de sa métaphore. Ces résultats sont congruents avec ceux de (Bouziane & Ziad, 2018b; Forrer & Wyant, 2015; Yarrow & Topping, 2001).

Alors même que la présente étude met plutôt l'accent sur la didactique du français langue première relativement à l'étude de la langue, à la production d'écrit et à l'articulation entre les deux, on constate que ce dispositif de coévaluation outillé par le numérique est, de fait, propice à une activité à la fois de scripteur et de lecteur, à l'instar d'expérimentations de didactisation de fanfiction en contexte scolaire (Brunel, 2018) ; toutefois les compétences

⁶⁴ Pour rappel, le score de coévaluateur est calculé par l'artéfact numérique en comparant, pour une production écrite donnée, l'inter-évaluation de l'élève avec l'hétéro-évaluation de l'enseignant, compétence par compétence.

cibles, opérationnalisées dans la grille de lecture/coévaluation co-construite, incitent ici le lecteur à focaliser son attention sur des caractéristiques linguistiques des écrits des pairs. A contrario, certains aspects créatifs ou esthétiques peuvent échapper à la nomenclature parfois restrictive de la grille de coévaluation. A cet égard, les élèves participants ont parfois regretté qu'une production écrite hors norme ne puisse pas être valorisée dans le cadre du dispositif de coévaluation numérique. Quoi qu'il en soit, ces expérimentations de didactisation de la production écrite à destination des pairs, outillées par le numérique, contribuent à une didactisation subséquente de la lecture. Compte tenu de la problématique représentée par les « non-lecteurs scolaires » aujourd'hui (Eugène, 2023), le dispositif expérimenté ici paraît donc revêtir un intérêt supplémentaire non négligeable.

1.8 Ajustements didactiques d'une itération à l'autre et effets sur la maîtrise des compétences cibles

Pour la seconde itération (T3/T4) du protocole en année N avec l'échantillon B, s'inspirant d'une activité mise en œuvre par l'enseignante A avec ses élèves de 3^e, l'enseignante B a spontanément opté pour un écrit court, en réponse à une question de compréhension de type DNB. Le nombre de critères d'évaluation a par ailleurs été abaissé à six. Ces deux choix constituent une régulation instrumentale (ajustement didactique) eu égard aux conflits instrumentaux didactiques rencontrés à T2 : avec un écrit long d'une part, et avec un trop grand nombre de critères d'évaluation d'autre part.

Au dire de l'enseignante, cet ajustement didactique s'avère fructueux pour ses élèves de 3^e, notamment parce que le format court permet de mieux les mobiliser. Par ailleurs, le temps consacré au projet est moindre qu'à T2 : la séance est moins chronophage, plus resserrée donc plus cohérente. Enfin, le développement des compétences et métacompétences afférentes est favorisé grâce au nombre restreint de critères dans le formulaire de coévaluation. L'analyse quantitative des données numériques concernées permet de confirmer cela. En effet, lorsqu'on compare les scores de scripteurs et de coévaluateurs des élèves de l'échantillon B d'une itération à l'autre (élève par élève et en moyenne, sur échantillons appariés), on constate que dans l'ensemble les scores de scripteurs ont progressé de manière significative entre la première et la seconde itération, tandis que les scores de coévaluateur sont demeurés presque stationnaires (Figure 75 et Figure 76).

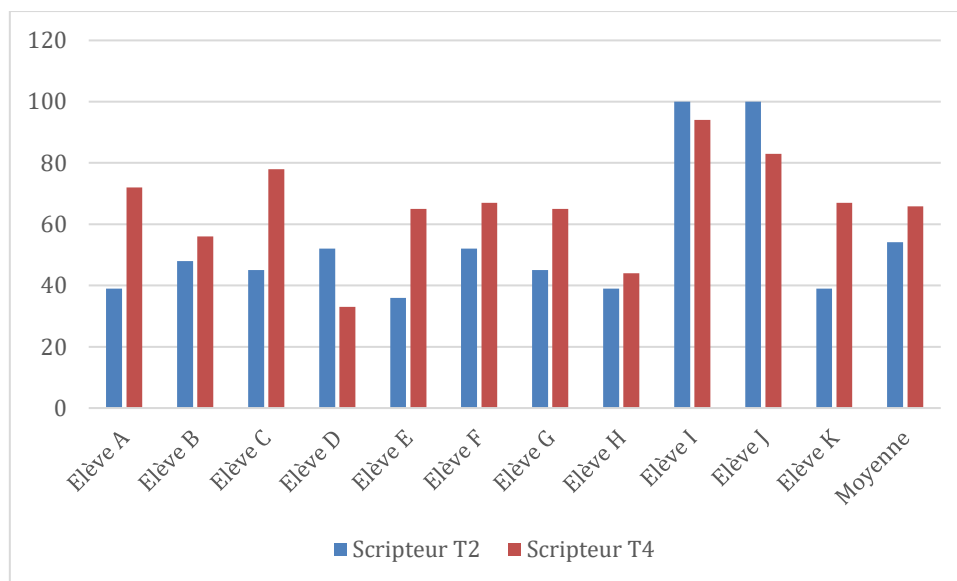


Figure 75 – Evolution des scores de scripteurs entre T2 et T4 pour l'échantillon B, après appariement (N = 11)

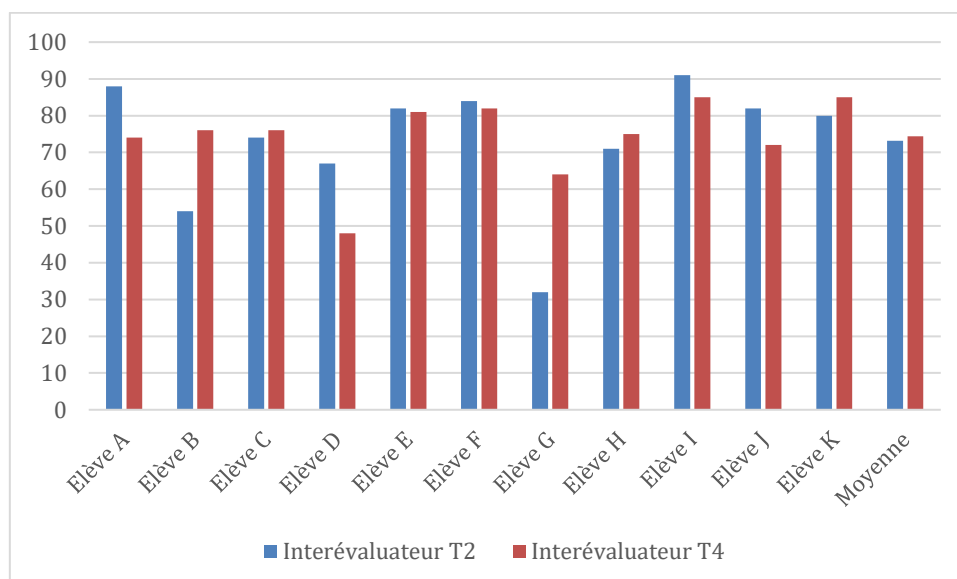


Figure 76 – Evolution des scores de coévaluateurs entre T2 et T4 pour l'échantillon B, après appariement (N = 11)

Ainsi, les tendances centrales rapportées ici montrent que les ajustements didactiques effectués lors de la seconde itération du protocole, en réponse aux conflits instrumentaux rencontrés lors de la première, ont eu un effet significatif sur l'amélioration des compétences des élèves en étude de la langue et production écrite en français langue première. En outre, bien que ces résultats témoignent d'une absence d'augmentation significative en termes de métacompétences entre T2 et T4, on peut supposer que le simple fait d'adopter, à deux reprises, cette posture métacognitive propre à la coévaluation, a contribué à l'amélioration constatée des scores de scripteurs, dans une logique de transfert des stratégies métacognitives d'une tâche complexe à l'autre (Efklides, 2006 ; Pintrich *et al.*, 2000 ; Romainville, 1995). En corroborant cette relation de causalité réciproque entre compétences linguistiques et scripturales d'une part,

et métacompétences coévaluatives d'autre part, nos résultats sont par ailleurs congruents avec ceux de recherches empiriques menées à plus grande échelle, à l'instar des travaux de (Hortigüela Alcalá *et al.*, 2019).

1.9 Artéfact numérique et prolongements possibles pour l'artéfact didactique

D'autres manières de mobiliser l'artéfact numérique Moodle Atelier permettraient de prolonger le dispositif pédagogique de coévaluation, dans le but d'approfondir l'articulation entre didactique de la langue et production d'écrit : soit par instrumentation, soit par instrumentalisation.

L'enseignante C signale qu'elle prolonge de manière manuscrite les apprentissages réalisés en salle informatique au moyen du dispositif pédagogique de coévaluation, notamment en offrant à ses élèves l'opportunité de retravailler leurs productions écrites en vue d'une nouvelle (hétéro-)évaluation, sommative cette fois. Par instrumentation, la catachrèse « boucle formative de rétroaction anonymisée » permettrait de mettre en œuvre via l'artéfact numérique une approche similaire, et de cibler les compétences à retravailler, de manière dynamique, pour aider un maximum d'élèves à manifester *in fine* une maîtrise satisfaisante des attendus didactiques.

Une autre approche consisterait, par instrumentation, à s'appuyer sur une fonctionnalité déjà présente dans Moodle Atelier et facilement mobilisable pour la coévaluation – bien qu'elle soit initialement prévue pour l'inter-évaluation. Il s'agit en l'occurrence de la fonctionnalité « Travaux publiés », qui permet là encore de mettre en exergue des compétences à retravailler, en les illustrant par des faits de langue issus de productions écrites d'élèves, anonymisées – sous réserve que l'élève concerné en soit d'accord. Cette pratique déjà de plus en plus courante en classe (Taous, 2020), peut ainsi être facilitée, presque systématisée par le truchement de l'artéfact numérique.

Les faits de langue ainsi étudiés sont issus d'écrits d'élèves, donc d'écrits auxquels leurs pairs peuvent s'identifier. Les questionnements et les commentaires qui surviennent naturellement à la faveur du dispositif pédagogique de coévaluation – *e.g.* face à un « dilemme docimologique » ou dans le cadre d'un écrit commentatif rédigé à destination d'un pair – sont autant de pistes pour organiser la suite des apprentissages en fonction des besoins réels des élèves.

Comme le déclare l'enseignante C à T6, le dispositif pédagonumérique est propice à une approche plus organique de la progression didactique en générale, singulièrement pour l'articulation entre étude de la langue et production d'écrit. A plus forte raison, les prolongements évoqués ici pour l'artéfact numérique, soit par instrumentalisation, soit par instrumentation, permettraient de faciliter, voire de systématiser ce renouvellement porteur pour la didactique du français langue première en cycle 4.

Ce qu'il faut retenir

Ce dispositif pédagonumérique semi-expérimental permet d'outiller la production écrite et la coévaluation pour cibler des compétences de cycle 4 en maîtrise de la langue et en production d'écrit. Nous avons vu qu'un certain nombre de conflits instrumentaux peuvent toutefois intervenir, gênant voire empêchant l'accès à l'artéfact didactique.

Pour commencer, un contresens pédagogique peut survenir à l'égard de la coévaluation : l'élève peut ne pas en percevoir l'utilité ni le sens ; ce conflit instrumental d'origine pédagogique peut entraîner un empêchement didactique. Nous avons vu également que des prérequis didactiques semblaient indispensables pour que l'élève puisse s'engager dans la tâche métacognitive qu'est la coévaluation, au risque d'un conflit instrumental pédagogique. A cet égard, les résultats de nos analyses qualitatives sur données textuelles et de nos analyses quantitatives sur données numériques convergeaient pour signifier une relation de causalité réciproque entre compétences linguistiques et scripturales d'une part, et métacompétence coévaluative d'autre part. Ce résultat est congruent avec les résultats d'études coévaluatives à plus grande échelle comme celle de (Hortigüela Alcalá *et al.*, 2019). Nous avons constaté par ailleurs que la coévaluation a posé des problèmes de discernement docimologique – à l'instar de l'inter-évaluation (Topping & Elliot, 2000). Les conflits instrumentaux peuvent se rapporter à l'artéfact numérique, plus précisément aux interfaces qui lui sont associées (*e.g.* clavier, correcticiel...), ce qui peut entraîner un double conflit instrumental, gênant l'accès à l'artéfact didactique à la fois pour l'élève scripteur et pour l'élève coévaluateur.

Toutefois, des processus de régulation instrumentale mobilisant une combinatoire d'artéfacts peuvent susciter des synergies régulatrices, à même de favoriser la motivation intrinsèque notamment, *e.g.* pour une intériorisation des consignes d'écriture (artéfact didactique). En fait, le dispositif pédagonumérique peut à la fois être source de conflits et de régulations

didactiques – régulations qui peuvent prendre appui sur la manière dont l’artéfact numérique est instrumentalisé. En d’autres termes, la catachrèse « coévaluation » induit des garde-fous susceptibles de réguler des conflits instrumentaux didactiques ou pédagogiques, y compris dans des cas limite – *e.g.* plagiat.

Des ajustements didactiques sont également susceptibles d’intervenir, d’une itération à l’autre, en réponse à un conflit instrumental précédemment identifié et à la faveur d’une évolution, par tâtonnements, des schèmes d’utilisation des enseignants. Nous avons vu également qu’un certain nombre de prolongements porteurs au plan didactique pourraient être mis en œuvre, soit par instrumentation, soit par instrumentalisation de l’artéfact numérique, en empruntant des pistes qui restent à explorer mais qui sont particulièrement encourageantes en raison d’une convergence entre construits théoriques et suggestions émanant directement des élèves participants et des enseignants partenaires.

Pour finir, ce dispositif ciblant la maîtrise de la langue et la production d’écrit s’est avéré, de fait, une didactisation de la lecture qui, à l’instar de certaines expérimentations avec la fanfiction en contexte scolaire par exemple (Brunel, 2018; Taous, 2020), peut constituer une piste intéressante en réponse à la problématique de la non-lecture scolaire (Eugène, 2023).

2 Coévaluation numérique et autodirection des apprentissages

Nous avons vu que nos résultats permettent de valider l’hypothèse HO 2, c’est-à-dire « La coévaluation numérique améliore l’autodirection des apprentissages ». A cet égard, en appui sur (Carré, 2010; Poisson *et al.*, 2015), nous avons pu établir que la coévaluation numérique contribue à l’augmentation de l’autorégulation des apprentissages, de la métacognition et du sentiment d’efficacité personnelle. Le présent chapitre propose de discuter ces résultats, en lien avec ceux afférents aux quatre autres hypothèses opérationnelles de la présente étude.

2.1 Coévaluation numérique et autorégulation des apprentissages

Concernant l'autorégulation des apprentissages, l'analyse quantitative de nos données numériques révèle que l'auto-efficacité perçue des élèves participants, relativement à l'autorégulation des apprentissages – c'est-à-dire leur score à la sous-échelle SEP(srl)⁶⁵ de l'échelle SEPE⁶⁶ –, tend à augmenter de manière significative au fil des itérations du protocole expérimental de coévaluation numérique. Or nous avons vu *supra*, avec (Carré, 2010; Ryan & Deci, 2000), que la motivation s'inscrit dans un « continuum de l'autodétermination ». A la lumière de cette typologie, on constate que le dispositif expérimental de coévaluation numérique est propice à une autodétermination accrue des apprentissages. Or pour rappel, l'autodétermination est une condition indispensable de l'autorégulation des apprentissages selon (Carré, 2010; Cosnefroy, 2010).

On peut objecter que, lorsque les élèves de « 6^e coop' » manifestent leur intérêt pour le dispositif dans la mesure où l'artéfact technique Moodle Atelier génère de facto une note supplémentaire, c'est-à-dire la note de coévaluateur – susceptible selon eux d'augmenter leur moyenne trimestrielle –, il s'agit d'une *motivation extrinsèque introjectée*, donc non-autodéterminée. Néanmoins, nous avons pu constater que le dispositif expérimental de coévaluation permet aux élèves de passer, d'une motivation extrinsèque contrôlée ou introjectée, à une motivation extrinsèque *identifiée*, en l'occurrence par adhésion aux valeurs coopératives sous-jacentes de l'artéfact pédagogique « coévaluation ». Nous avons vu en effet que les élèves apprécient le « renversement pédagogique », source de ludification, qui leur permet de « jouer au prof ».

A fortiori, cette activité de coévaluation consiste en quelque sorte pour les élèves en un « jeu sérieux » (Ouherrou, 2022) dans la mesure où ils insistent souvent sur le défi que représente pour eux le discernement docimologique nécessaire à une évaluation « juste » – au sens de « précise », mais aussi selon l'acception éthique du terme. De même, un nombre significatif d'élèves souligne le souci qu'ils ont de rédiger à destination des pairs qu'ils évaluent un écrit commentatif pertinent, intelligible, argumenté, constructif et bienveillant. Une

⁶⁵ SEP(srl) : Sentiment d'efficacité personnelle pour l'autorégulation des apprentissages (communément appelée « *self-regulated learning* » dans la littérature anglophone).

⁶⁶ SEPE : Sentiment d'efficacité personnelles des enfants (et des adolescents).

telle préoccupation correspond parfois à une motivation extrinsèque intégrée – la forme de motivation extrinsèque la plus autodéterminée selon (Ryan & Deci, 2000) ; en effet, on se trouve alors en présence d'une adhésion profonde, librement déterminée, à l'artéfact pédagogique, qui confine à la motivation intrinsèque dans la mesure où la gratification recherchée par l'élève est inhérente à la tâche elle-même, et repose *in fine* sur la motivation altruiste d'aider un pair à progresser, en lui fournissant la meilleure coévaluation et le meilleur feedback possibles. Ce type de motivation altruiste, manifeste dans certains verbatim d'élèves, correspond donc à une motivation intrinsèque ou, à minima, à une motivation extrinsèque intégrée.

Une autre situation dans laquelle nous avons pu constater une évolution manifeste, d'une motivation extrinsèque contrôlée ou introjectée, vers une motivation autodéterminée, consiste en l'intériorisation d'une consigne d'écriture, initialement mal comprise ou jugée absurde, à la faveur d'exemples concrets, sources d'inspiration voire d'émulation, rencontrés dans des écrits de pairs coévalués. Le cas de Marie et de son rapport à la consigne d'écriture « métaphore » illustre bien ce phénomène. Ce cas fait l'objet d'une attention et d'une documentation spécifiques *infra*, dans la mesure où il a servi de support à une partie des échanges avec l'enseignante C à T6 ; toutefois les données textuelles recueillies auprès des autres élèves montrent qu'il ne s'agit pas d'un cas isolé. Ici, il ne s'agit pas simplement d'une amotivation évoluant vers une motivation contrôlée ou une motivation introjectée, mais bien d'une évolution vers l'autodétermination, c'est-à-dire en faveur d'une motivation identifiée voire intégrée, puisqu'à la faveur du dispositif de coévaluation l'élève initialement récalcitrant ou en difficulté (conflit instrumental didactique) parvient à faire siennes les valeurs induites à la fois par l'artéfact pédagogique – source d'ouverture sur les écrits des pairs – et par l'artéfact didactique, en l'occurrence la consigne d'écriture comme contrainte paradoxalement source de créativité et de plaisir esthétique.

En somme, le dispositif expérimental de coévaluation numérique favorise effectivement l'autorégulation des apprentissages, en permettant aux élèves de progresser vers une motivation davantage autodéterminée, selon la typologie motivationnelle de l'autodétermination établie par (Ryan & Deci, 2000) et conformément à l'articulation entre motivation et volition théorisée par (Cosnefroy, 2010; Zimmerman, 2008), ce qui tend à expliquer l'augmentation constatée du score obtenu par l'ensemble des élèves participants à l'échelle SEP(srl) mesurant l'auto-efficacité perçue pour l'autorégulation des apprentissages, pendant l'ensemble de la période expérimentale.

2.2 Coévaluation numérique, métacognition et sentiment d'efficacité personnelle

Nous avons vu *supra* que la motivation joue un rôle central dans l'autorégulation des apprentissages, or elle est fortement liée à la notion de métacognition (Berger & Büchel, 2012; Zimmerman, 1995). A ce titre, l'analyse quantitative des données numériques révèle une augmentation du score de coévaluateur de l'ensemble des élèves participants, en particulier à T4, ce qui indique une amélioration des compétences métacognitives des élèves participants.

Rappelons que (Berger & Büchel, 2012, p. 96), en appui sur (Efklides, 2006; Pintrich *et al.*, 2000), distinguent trois composantes de la métacognition : les connaissances métacognitives, les expériences métacognitives et les stratégies métacognitives.

Nous avons vu que les connaissances métacognitives ou « métaconnaissances » sont stockées dans la mémoire à long-terme de l'apprenant et peuvent *in fine* se décliner en deux catégories, selon qu'elles relèvent d'un savoir ou d'un savoir-faire⁶⁷. Or la présente recherche montre que certains savoirs et/ou savoir-faire en lien avec l'artéfact didactique peuvent constituer un prérequis pour bénéficier des apprentissages, à la fois en tant que scripteur – c'est-à-dire pour monter en compétences, relativement aux attendus didactiques en maîtrise de la langue et production d'écrit en français langue première – et en tant que coévaluateur, c'est-à-dire pour « monter en métacompetences ».

Par ailleurs, selon (Efklides, 2001, cité par Berger & Büchel, 2012, p. 96), les expériences métacognitives constituent « des sentiments et jugements subjectifs » en lien avec le(s) processus cognitif(s) en cours lors de la réalisation d'une activité d'apprentissage. Ces expériences surviennent quand l'élève « prend conscience de ses processus de pensée actuels » et leur rôle principal est de « guider les processus de pensée d'un point de vue subjectif et du point de vue des connaissances spécifiques à la tâche, ceci afin d'influencer les stratégies métacognitives ». Or le dispositif expérimental de coévaluation numérique permet d'articuler maîtrise de la langue, production d'écrit et lecture. Les processus cognitifs et les prérequis didactiques mobilisés pour effectuer une tâche complexe en production d'écrit ciblant la maîtrise de la langue sont donc complétés, dans un second temps, par des processus cognitifs

⁶⁷ Selon (Büchel, 1991; VandenBos, 2007), on peut articuler d'emblée savoir-faire et savoir-être, dans la mesure où on peut considérer que toute stratégie est intrinsèquement procédurale et inclut ainsi une connaissance des circonstances propices à son application.

en lien avec une didactisation de la lecture. Outillée par le numérique, cette activité de « lecture métacognitive » est d'autant plus active qu'elle s'inscrit dans le cadre d'une pédagogie constructiviste, plaçant l'élève dans une posture d'élève-chercheur (Barth, 2013, 2015), avec un étayage didactique fondé sur une grille de coévaluation co-construite en amont par l'enseignante partenaire et ses élèves. Enfin, un troisième « étage » métacognitif permet à l'élève de réaliser en quelque sorte la synthèse des activités d'écriture et de lecture préalables, en formulant à destination du pair coévalué un écrit commentatif argumenté. On peut ainsi supposer qu'en dehors des conflits instrumentaux susceptibles d'enrayer sa genèse instrumentale – dont nous avons vu par ailleurs qu'ils s'accompagnent le plus souvent de processus régulateurs – le dispositif expérimental de coévaluation numérique constitue une authentique « expérience métacognitive » au sens de (Berger & Büchel, 2012; Efklides, 2001).

En outre, (Berger & Büchel, 2012) précisent que les jugements subjectifs formulés par l'apprenant dans le cadre de ses expériences métacognitives concernent la difficulté estimée de la tâche ou la confiance relative dans l'exactitude d'une solution que l'élève a trouvée pour celle-ci. Ces expériences peuvent aussi prendre « la forme de sentiments tels que le sentiment de familiarité avec la tâche ». Nous avons donc postulé que le sentiment d'efficacité personnelle est lui aussi amené à jouer un rôle important dans ces expériences métacognitives. Or lorsque pour l'ensemble des élèves participants nous avons effectué une analyse de corrélation au moyen du test de Spearman sur les scores de coévaluateurs et les scores individuels moyens obtenus à l'échelle SEPE (exprimés en pourcentage et sur échantillons appariés), bien qu'une corrélation significative n'ait pas pu être établie à T4, nous avons pu établir une corrélation positive à T2 ($p = 0.031$, $r = 0.402$, $N = 29$), représentée dans la Figure 77. Pour l'ensemble des élèves participants, on peut ainsi dire qu'à T2 l'augmentation des scores de coévaluateurs coïncide avec une augmentation des scores individuels moyens à l'échelle SEPE. Bien que leur représentativité soit limitée, ces résultats sont congruents avec les thèses de (Berger & Büchel, 2012; Efklides, 2001) relativement aux interactions entre métacognition et auto-efficacité perçue.

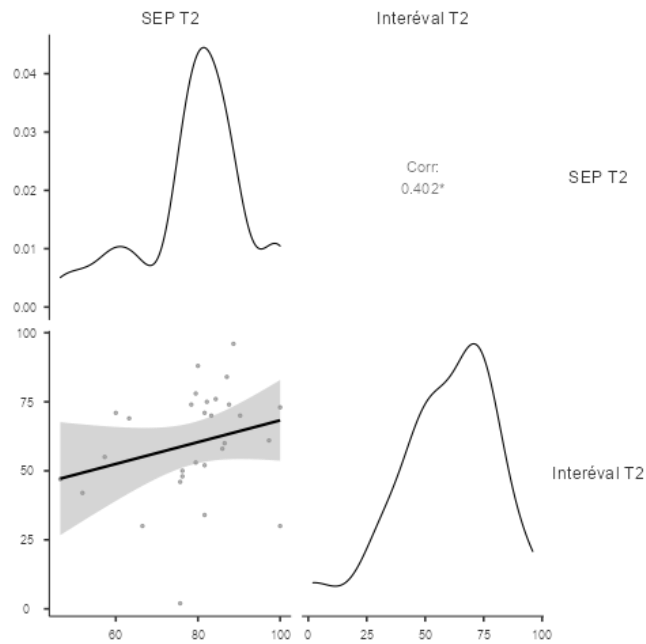


Figure 77 – Matrice de corrélation représentant les résultats du test de Spearman pratiqué sur les scores de coévaluateurs et les scores individuels moyens à l'échelle SEPE, échantillon SE, T2, N = 29 (en %, après appariement)

Plus largement, (Zimmerman, 1995) met l'accent sur l'interdépendance entre métacognition et motivation pour favoriser les apprentissages des élèves. Notamment parce que « des croyances motivationnelles défavorables sont liées à un moindre exercice de stratégies métacognitives ainsi qu'à des jugements métacognitifs moins adaptés » (Berger & Büchel, 2012, p. 117). Ainsi, pour mieux analyser les interactions entre motivation et métacognition, (Zimmerman, 1995) propose explicitement d'opérationnaliser cette notion de « croyance motivationnelle » au moyen du sentiment d'efficacité personnelle de (Bandura, 2007). Or nous avons pu valider l'hypothèse selon laquelle le dispositif expérimental de coévaluation numérique contribue à l'augmentation du sentiment d'efficacité personnelle des élèves participants. Par ailleurs, sur la base de leur vécu expérimental en lien avec la présente recherche collaborative et en se fondant sur leur intuition professionnelle, la plupart des enseignantes partenaires perçoivent explicitement l'intérêt d'une pratique régulière de la coévaluation numérique, pour favoriser un climat de classe coopératif, propice à la coévaluation et au développement des métacompétences.

Pour opérationnaliser plus avant ces relations entre motivation, sentiment d'efficacité personnelle et engagement (méta-)cognitif, (Gauthier *et al.*, 2007) soulignent l'importance de l'enseignement explicite, qui pourrait constituer un modèle pertinent pour encourager le développement de croyances motivationnelles propices à l'acquisition des savoirs et des savoir-faire. Dans ses travaux, (Schunk, 1982) expérimentait déjà ce modèle d'enseignement explicite

en mettant en lumière le rôle bénéfique du feedback pour les apprentissages des élèves : un dispositif qui se révèle fructueux pour le « développement du sentiment d'efficacité personnelle chez des élèves rencontrant des difficultés d'apprentissage » (Berger & Büchel, 2012, p. 119). Or dans le cadre du dispositif expérimental de coévaluation, nous avons pu constater le rôle central joué par les écrits commentatifs d'élèves, qui semblent faire l'objet d'une attention particulière de leur part, en termes de production – quelle que soit leur maîtrise des compétences didactiques ciblées par le dispositif. Cette production écrite d'un écrit commentatif à destination de leurs pairs paraît en effet être considérée avec sérieux – peut-être même vue comme un défi. A la lumière des verbatim d'élèves rappelés dans le Tableau 29, on constate effectivement que la formulation d'écrits commentatifs argumentés pose problème à des élèves aux profils variés en termes de performances de scripteurs et de coévaluateurs – ces deux indicateurs étant d'ailleurs rarement corrélés en ce qui concerne l'échantillon restreint présenté ici. D'autre part, on constate que les trois derniers élèves obtiennent des scores de coévaluateurs élevés, alors que le verbatim afférent exprime un faible sentiment d'efficacité en la matière.

Verbatim d'élèves	Moyenne coévaluateur (T2-T4, en %)	Moyenne scripteur (T2-T4, en %)
Car il faut donner de bonnes explications et trouver la bonne justification.	45	25
Car il est difficile de mettre des commentaires mais sinon je trouve ça plutôt facile.	50	69
Il faut savoir trouver les bons mots.	51	42
J'ai trouvé difficile le fait de justifier les choix de la coévaluation.	52	28
C'est assez dur de trouver des arguments pour lui expliquer ce qui ne va pas et comment l'améliorer et lui donner une compétence globale pour son travail, pour lui permettre de comprendre et d'améliorer ses textes.	72	52
C'est dur de trouver les mots qu'il faut pour évaluer correctement un travail.	72	94

Car il y a des choses que je n'arrive pas à dire mais que j'ai réussi à répondre.	73	42
---	----	----

Tableau 29 – Engagement (méta-)cognitif dans la formulation d'écrits commentatifs argumentés à destination des pairs coévalués

En ce qui concerne la réception des écrits commentatifs, nous avons vu qu'ils ne sont pas toujours jugés suffisamment explicites par les pairs coévalués. A l'occasion des focus groupes, les élèves participants ont d'ailleurs insisté sur le fait que tout écrit commentatif doit être anonyme, constructif (*e.g.* pas un commentaire laconique et stérile de type « c'est nul ») et argumenté – pour appuyer à la fois la légitimité de la coévaluation, son acceptabilité, et pour permettre à l'élève évalué d'effectuer des progrès réels. Ainsi, le feedback entre pairs peut faciliter la métacognition dans la perspective d'un enseignement explicite, comme le préconisent (Gauthier *et al.*, 2007; Schunk, 1982), en vue d'une augmentation du sentiment d'efficacité personnelle des élèves (Bandura, 2007; Berger & Büchel, 2012) ; il n'en demeure pas moins que ce dernier peut également être bridé, en production, par des croyances motivationnelles défavorables et des difficultés liées aux métacompétences à mobiliser pour argumenter. En retour, le feedback entre pairs peut poser des problèmes d'acceptabilité, en réception, de la part des pairs évalués, dans la mesure où s'il est insuffisamment argumenté, ces derniers peuvent remettre sa légitimité en cause. En émission comme en réception, un tel feedback entre pairs gagne cependant à être anonymisé en double aveugle – seul l'enseignant sait qui a écrit quoi et pour qui. Cette anonymisation, qui est l'apanage de l'artéfact numérique, permet d'éviter – autrement dit de réguler – des biais sociocognitifs qui constitueraient un potentiel conflit instrumental pédagogique, gênant l'accès à l'artéfact didactique – et partant, aux métacompétences afférentes.

(Berger & Büchel, 2012) indiquent par ailleurs que l'articulation entre la métacognition et le sentiment d'efficacité personnelle des élèves peut également bénéficier du « modelage », c'est-à-dire lorsque l'enseignant lui-même se donne en exemple dans la réalisation d'une tâche, explicitant ses stratégies métacognitives pour que les élèves puissent se les approprier. Or nous avons vu que le dispositif expérimental de coévaluation a suscité chez l'enseignante C la volonté de pratiquer à plusieurs reprises cette forme de « modelage » avec ses élèves. Les travaux de (Schunk & Ertmer, 2000) ont montré que cette pratique peut être d'autant plus profitable pour les apprenants – notamment ceux qui sont le plus en difficulté – s'ils modèlent leurs stratégies métacognitives d'après celles de leurs pairs qui sont davantage en réussite ; à condition toutefois de demeurer dans leur zone de développement proche (Vygotski, 1985), *i.e.*

en s'inspirant du travail d'élèves dont le degré de maîtrise est supérieur au leur, sans en être trop éloigné (Reynaud, 2022).

Ce qu'il faut retenir

En somme, le dispositif de coévaluation numérique individualisée qui fait l'objet de la présente étude permet de favoriser l'autorégulation des apprentissages. La relation dialectique dont parle (Carré, 2010, p. 133), « entre facteurs motivationnels et métacognitifs de l'autodirection » des apprentissages, se trouve ici empiriquement confirmée, à la fois par les résultats de nos analyses quantitatives sur données numériques et par les résultats de nos analyses qualitatives sur données textuelles. De même, cette recherche permet de confirmer l'étroite association entre l'autodétermination des apprentissages (Deci & Ryan, 2008), l'autorégulation des apprentissages (Cosnefroy, 2010; Pintrich, 2000; Zimmerman, 2008), et le sentiment d'efficacité personnelle (SEP) théorisé par (Bandura, 2007). A cet égard, nous avons vu que le SEP, « coefficient » de l'autodirection des apprentissages selon (Carré, 2010, p. 156), joue ici un rôle transversal : en effet, avec la mise en œuvre de ce dispositif expérimental de coévaluation numérique, nous avons pu constater empiriquement les relations que ce concept entretient avec la motivation, la métacognition et l'autodétermination. De surcroît, par l'entremise des sous-échelles SEP(srl), SEP(coop) et SEP(scol), nous avons pu mettre en évidence les implications du SEP pour le dispositif, respectivement en termes d'autorégulation des apprentissages, relativement aux apprentissages coopératifs et au plan didactique.

3 La coévaluation numérique est-elle coopérative ?

Les résultats des analyses quantitatives pratiquées sur données textuelles et numériques permettent de valider l'hypothèse HO 3, *i.e.* « La coévaluation numérique améliore les apprentissages coopératifs des élèves ». A la faveur d'un retour réflexif sur le cadre théorique afférent, et selon une approche transversale visant à croiser les résultats liés à nos quatre hypothèses opérationnelles, ce chapitre propose d'approfondir la discussion.

3.1 Coopération entre élèves et enseignants

Le dispositif s'avère coopératif par nature, dans la mesure où il implique une coopération entre l'enseignant et ses élèves, qui « partagent l'évaluation ». En effet, la comparaison que permet l'artéfact numérique entre l'hétéro-évaluation de l'enseignant et l'inter-évaluation de l'élève constitue une manifestation concrète de cette coopération entre eux : ils travaillent ainsi conjointement à la réussite de la situation d'enseignement/apprentissage. Nous avons vu toutefois que certains élèves ne perçoivent pas bien les enjeux métacognitifs liés à la coévaluation, et semblent voir l'acte évaluatif comme quelque chose d'arbitraire. Il s'agit donc d'un contresens pédagogique à l'égard de l'artéfact « coévaluation », mais aussi potentiellement d'un « contresens didactique » relativement à la grille de coévaluation co-construite et aux compétences cibles qu'on se propose d'évaluer par son entremise – à fortiori par le truchement rationalisé d'une échelle de Likert à quatre points, qui oblige donc à trancher, et dont on a vu qu'elle génère des données non pas simplement ordinales mais « à intervalle ». Cette vision qui oppose hétéro- et inter-évaluation (Tableau 30) constitue un conflit instrumental d'origine pédagogique et didactique qui peut entraver la dimension coopérative de la situation d'enseignement/apprentissage.

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim)	Moyenne scripteur	Moyenne coéval.
Plutôt difficile	On n'a pas les mêmes points de vue que les profs.	67	50
Facile et difficile à la fois	Car je peux aimer le travail et bien l'évaluer et le prof qui va l'évaluer peut moins aimer le travail et mettre une note opposée à la mienne.	60	83

Tableau 30 – Verbatim d'élève illustrant la vision conflictuelle entre hétéro- et inter-évaluation

Toutefois, y compris dans le contexte d'une classe pratiquant déjà fortement les pédagogies coopératives – *i.e.* la « classe coop' » C –, nous avons vu que le dispositif contribue à un approfondissement des pratiques coopératives, en faveur notamment d'une objectivation accrue des pratiques évaluatives et d'un affinement de la réflexion docimologique de la part de l'enseignante partenaire, nourris par les coévaluations des élèves – mais aussi par leurs questions et leurs commentaires avant, pendant et après l'activité de coévaluation stricto sensu.

Nous avons vu également que ces éléments sont de nature à alimenter la progression didactique de manière plus organique, singulièrement pour l’articulation entre étude de la langue et production d’écrit, en lien avec les questionnements spontanés que le dispositif expérimental de coévaluation contribue à faire émerger parmi ce collectif apprenant (Lescouarch, 2021) – en réaction à d’authentiques écrits d’élèves, auxquels leurs pairs peuvent s’identifier.

3.2 Coopération entre élèves

On peut se demander avec (Zarifian, 2009, pp. 54-55) si le dispositif expérimental de coévaluation numérique que nous proposons est réellement coopératif :

« Les applications informatiques : elles sont souvent utilisées comme un grand économiseur de coopération. On utilise et on entre des informations selon un codage et un formatage déterminés dans un logiciel. L’application peut ainsi coordonner, sinon les activités elles-mêmes (bien que certaines applications puissent conditionner et déclencher l’activité d’une personne), du moins leur interdépendance informationnelle, sans que les personnes ne coopèrent, voire souvent : sans qu’elles ne se connaissent. (...) Pour coopérer, il faut partager le sens de ce que l’on va faire ensemble. Le sens, c’est à la fois l’orientation, le vers quoi, mais aussi la valeur, l’importance subjective discriminante que chacun donne à cet agir en commun. »

Certes, les élèves ne sont pas amenés à coopérer en face-à-face ; toutefois la coopération n’en semble pas moins effective, dans la mesure où ils coévaluent effectivement les écrits de plusieurs de leurs pairs. Au-delà de la grille de coévaluation critériée, qui constitue déjà, en soi, une première forme de feedback, les élèves évaluateurs fournissent également à leurs pairs un écrit commentatif. Nous avons vu que ce dernier n’est pas toujours formulé de manière suffisamment explicite ; pour autant il fait globalement l’objet d’un engagement significatif de la part des élèves, à la fois en réception et en émission. La dimension coopérative du dispositif pédagogique numérique est par ailleurs mise en avant par les élèves, au sens où elle leur permet de « corriger les copies des camarades » en « faisant le prof », ce qui est « cool ». Loin d’être anecdotiques, ces éléments de langage – qui sont très prégnants dans les données textuelles recueillies auprès des élèves (Annexes 1.5.4, Tableau 52) – manifestent une ludification latente de la métacognition, induite par le dispositif de coévaluation numérique. De surcroît, les élèves

insistent sur l'opportunité qui leur est ainsi offerte de découvrir les productions écrites de leurs pairs, sources d'inspiration voire de motivation intrinsèque pour intérioriser une consigne d'écriture ou une compétence cible, grâce à des exemples auxquels ils peuvent s'identifier.

3.3 Relation d'aide

Relativement aux quatre formes identifiées par (Connac, 2017b) pour la coopération entre élèves, on peut parler ici de relation d'aide dans la mesure où, qu'il soit « expert » ou non, relativement aux compétences cibles, l'élève évaluateur est amené par l'artéfact pédagogique à adopter, de facto, une posture métacognitive vis-à-vis de l'écrit du pair évalué. A fortiori, nos résultats montrent qu'on se situe effectivement ici dans des configurations correspondant à ce que (Crahay *et al.*, 2001, p. 135) nomment des « tactiques rétroactives » d'aide, *i.e.* lorsque l'aidé a d'abord cherché à résoudre la tâche par lui-même et que l'aidant lui fournit une explication à postériori, en « réagissant aux erreurs de [l'aidé] et/ou à ses questions ». En effet, les scénarisations mises en œuvre dans le cadre du présent protocole expérimental amènent chaque élève à être confronté dans un premier temps à une situation-problème pour produire son écrit. En outre, lorsque l'élève reçoit le feedback de ses pairs et de l'enseignante (*i.e.* formulaires de coévaluation et écrits commentatifs), ce dernier intervient effectivement à postériori. De fait on peut dire, dans le prolongement des travaux de (Barth, 2013), que l'aidé comme l'aidant adoptent tous deux une posture constructiviste d'élève chercheur. En outre, il y a réversibilité de la relation d'aide dans la mesure où le pair évalué – l'aidé – est simultanément un aidant qui évalue les écrits des autres.

3.4 Souci de formuler un écrit commentatif adapté

La simultanéité qui caractérise ici la réversibilité de la relation d'aide est une spécificité du numérique, liée à l'hybridité d'un dispositif qui s'avère à la fois synchrone et asynchrone, présentiel et distanciel – du fait du travail dans une seule et même salle informatique certes, mais par ordinateurs interposés. Pour autant, l'informatique ne semble pas « faire écran » à la coopération, au contraire : les élèves signalent l'importance que revêt pour eux cette dimension altruiste du dispositif, particulièrement manifestée par le souci de fournir un écrit commentatif de qualité au pair évalué/aidé (Tableau 31). Les moyennes de scripteur et de coévaluateur des auteurs de ces verbatim montrent que ces élèves, qui manifestent un souci fort concernant la formulation d'un écrit commentatif adapté, peuvent avoir des profils très divers : de l'élève en

difficulté à l'élève expert, avec différentes combinaisons entre compétences de scripteur et métacompétences de coévaluateur – ce qui montre combien le dispositif amène structurellement tout élève à se placer, de facto, sinon dans une position d'expert, du moins dans une position d'aidant, et ce quel que soit son profil.

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim)	Moy. Scripteur (%)	Moy. Coéval. (%)
Facile	Car il est difficile de mettre des commentaires mais sinon je trouve ça plutôt facile	69	50
Facile et difficile à la fois	Il faut savoir trouver les bons mots	42	51
Facile et difficile à la fois	Car il y a des choses que je n'arrive pas à dire mais que j'ai réussi à répondre	42	73
Facile et difficile à la fois	J'ai trouvé difficile le fait de justifier les choix de la coévaluation	28	52
Plutôt facile	C'est assez dur de trouver des arguments pour lui expliquer ce qui ne va pas et comment l'améliorer et lui donner une compétence globale pour son travail, pour lui permettre de comprendre et d'améliorer ses textes.	52	72
Plutôt facile	Car il faut donner de bonnes explications et trouver la bonne justification	25	45
Facile et difficile à la fois	C'est dur de trouver les mots qu'il faut pour évaluer correctement un travail	94	72

Tableau 31 – Verbatim d'élèves illustrant le souci d'aider le pair évalué avec un écrit commentatif adapté

3.5 Travail de groupe et entraide

A priori ce dispositif de coévaluation individualisée n'est pas propice à ce que (Connac, 2017b) appelle « travail de groupe », ou à ce qu'il nomme « entraide » ; pourtant une situation imprévue nous a permis de constater que cela peut être le cas, quoique sous certaines

conditions. Pour définir l'entraide, (Connac, 2017b, p. 27) évoque « une interaction entre plusieurs personnes, conjointement bloquées face à une même difficulté. Ces personnes se réunissent de manière informelle, c'est-à-dire à leur initiative, pour tenter de la résoudre ensemble. Elles s'organisent librement ». Il précise que tous les apprenants concernés se trouvent alors au « même niveau de compétence quant à l'objectif qui les unit : on se trouve donc dans une relation symétrique ». Pour rappel, lors de la séance expérimentale à T1/T2 avec la classe A' (4^e), suite à une panne affectant plusieurs ordinateurs en salle informatique, trois élèves sont allés s'installer au centre de documentation et d'information (CDI) du collège, pour effectuer leurs coévaluations. Or à la faveur de cette reconfiguration imprévue, les trois élèves en question se sont mis à coopérer spontanément, se lançant dans des « débats docimologiques » pour déterminer la juste évaluation pour tel ou tel cas jugé litigieux. Les échanges étaient très « horizontaux », sans que personne ne prenne l'ascendant sur les autres, par exemple en raison d'une plus grande maîtrise des compétences cibles. Il ne s'agissait pas non plus d'une réelle demande d'aide, mais « plutôt [de] se rassurer, car [l'élève] n'osait simplement pas le faire de lui-même ». A cet égard il paraît intéressant de préciser que l'élève ayant initié ces échanges a obtenu un très bon score de coévaluateur (75%) : il semble donc avoir bénéficié de ces moments de coopération avec ses pairs. En somme, cette reconfiguration imprévue de l'espace d'apprentissage met en avant une dimension groupale spontanée, informelle et horizontale qui correspond bien à la définition donnée plus haut, et qui montre combien l'hybridité du dispositif est à ajustement variable : les écrans d'ordinateurs n'empêchent pas de se pencher vers le voisin pour initier des échanges complémentaires en face-à-face. A plus forte raison, avec des outils plus mobiles de type tablette, on pourrait imaginer des scénarios pour lesquels cette dimension fait d'emblée partie du dispositif.

Ce qu'il faut retenir

En résumé, le dispositif expérimental de coévaluation numérique permet de mettre concrètement en œuvre les pédagogies coopératives au sein des classes partenaires, ce dont attestent les résultats des analyses quantitatives pratiquées sur données textuelles et numériques. En outre, il s'agit d'une authentique coopération au sens de (Zarifian, 2009), avec un vouloir commun, co-construit. Nous avons d'ailleurs vu que cette coopération ne concerne pas seulement les élèves entre eux, mais aussi les enseignants et les élèves – avec en classe coopérative un approfondissement des schèmes d'enseignement déjà présents à cet égard. Nous avons également vu que la coopération entre élèves peut prendre la forme d'une

relation d'aide – avec un accent particulier mis sur le souci de formuler un écrit commentatif adapté, constructif, argumenté, pour aider les pairs coévalués à progresser. En outre, le dispositif s'avère propice au travail de groupe et à l'entraide. S'il n'inclue pas le tutorat, nous avons vu que le dispositif en permet l'instrumentation au moyen d'une fonctionnalité spécifique de l'artéfact numérique : les « Travaux publiés », pour des formes hybrides de tutorat, instrumentées par l'artéfact numérique – ou issues d'une instrumentalisation de celui-ci, et qui reste à inventer (*i.e.* catachrèse). Ces formes d'apprentissage coopératif constituent autant d'opportunités pour entretenir une dynamique formative efficace, propice au retravail des écrits, à l'approfondissement des compétences, et dans le cadre duquel le feedback (*i.e.* formulaires d'évaluation et écrits commentatifs) occupe une place centrale. A ce titre, l'anonymisation en double aveugle s'avère déterminante pour prévenir les risques identifiés par (Dekhinet *et al.*, 2008; Dochy *et al.*, 1999; Topping, 1998) en matière de biais sociocognitifs ou socioaffectifs. Dans le prolongement des pistes d'instrumentalisation précédemment évoquées, nous avons vu enfin que le dispositif pourrait évoluer pour constituer une véritable boucle de rétroaction formative anonymisée, à la faveur d'une catachrèse qui a d'ailleurs été suggérée par les élèves eux-mêmes, comme piste de régulation pédagonumérique en réponse à un conflit instrumental didactique. Une telle boucle évaluative constituerait une instrumentation numérique intéressante du concept développé en la matière par (Connac, 2017a; Connac & Cieutat, 2021; Reynaud, 2022).

4 Coévaluation numérique et « dialectique instrumentale »

Pour rappel, les analyses inférentielles pratiquées afin de quantifier le degré de conflictualité instrumentale, ressenti par les élèves participants vis-à-vis du dispositif expérimental de coévaluation numérique, n'ont généré aucun résultat significatif. Pour l'ensemble de la période expérimentale, l'analyse des données descriptives permet cependant d'identifier une tendance centrale selon laquelle, en moyenne, les élèves participants (super-échantillon SE, N = 57) ont répondu « La coévaluation avec Moodle Atelier m'a paru facile » au questionnaire – *i.e.* un score moyen de 4 sur 5, où 5 correspond au degré de conflictualité minimal dans l'échelle de Likert. Ce résultat laisse entendre une relative absence de conflit instrumental lors des situations d'enseignement/apprentissage mises en œuvre avec le dispositif expérimental de coévaluation. De surcroit, concernant l'échantillon C, cette tendance centrale

se trouve corroborée par les résultats issus des retours écrits sur expérience et des focus groupes. De fait, c'est surtout la combinaison des résultats issus des analyses qualitatives, pratiquées sur les données textuelles recueillies auprès des élèves participants et des enseignantes partenaires, qui nous permet de valider l'hypothèse HO 4, *i.e.* « La coévaluation numérique génère des conflits instrumentaux, mais permet de les réguler ».

Nous avons vu que, du point de vue des élèves, ces conflits instrumentaux se rapportent toujours à une combinatoire impliquant plusieurs des trois artéfacts – didactique, pédagogique et technique – qui y sont enchâssés (Marquet, 2005a). Du point de vue des enseignantes partenaires – *i.e.* d'après l'analyse qualitative des données textuelles recueillies auprès d'elles –, ceux-ci se rapportent aux artéfacts didactique et technique du dispositif expérimental, ainsi qu'à une combinatoire pédagogique et technique. En outre, dans le prolongement des travaux de (Marquet, 2003, 2005b, 2011; Marquet & Coulibaly, 2011), nous avons pu proposer une typologie, selon que les conflits instrumentaux identifiés entraînent un retard pour la situation d'enseignement/apprentissage, ou bien un renoncement partiel – voire total – à l'activité. Toutefois, nous avons constaté que les conflits instrumentaux identifiés s'accompagnent, le plus souvent, de processus de régulation instrumentale qui permettent d'en pallier partiellement ou totalement les effets sur la genèse instrumentale – tantôt simultanément, tantôt en asynchronie (Zeller & Marquet, 2021b). A cet égard, il s'est révélé nécessaire de distinguer dans l'analyse quels artéfacts – ou combinaisons d'artéfacts – sont à la source d'un conflit instrumental (« artéfacts conflictuels »), quels artéfacts s'en trouvent affectés (« artéfacts affectés »), enfin quels artéfacts sont susceptibles de réguler le conflit instrumental identifié, *i.e.* les « artéfacts régulateurs » – et, singulièrement, les « fonctionnalités régulatrices » qui sont l'apanage de l'artéfact numérique.

4.1 Combinatoires conflictuelles et synergies régulatrices

Du point de vue des élèves, parmi les combinaisons conflictuelles identifiées, on peut citer des inhibitions socioaffectives, une perception arbitraire de l'évaluation, la difficulté à formuler un écrit commentatif adapté, des mésusages (ou non-usages) des interfaces liées à l'artéfact numérique et un manque de prérequis didactiques. Du fait de l'emboîtement qui caractérise les trois artéfacts didactique, pédagogique et numérique dans la genèse instrumentale en EIAH, ces combinaisons conflictuelles sont caractérisées par des interactions complexes, avec parfois des relations de causalité réciproque ou des effets rétroactifs. Cette dimension dialectique concerne également les « artéfacts régulateurs ». De fait, lorsqu'on

distingue entre artéfacts conflictuels, affectés et régulateurs, on s'aperçoit que certains conflits instrumentaux peuvent affecter plusieurs artéfacts, *e.g.* la combinatoire conflictuelle en lien avec les mésusages (ou non-usages) des interfaces numériques. Nous avons d'ailleurs observé que ces combinaisons conflictuelles peuvent concerner à la fois le pair coévaluateur et le pair coévalué – quoique pour des artéfacts différents.

A titre d'exemple, on peut également citer une perception arbitraire de l'acte évaluatif en général. En effet, nous avons vu qu'une combinatoire conflictuelle d'origine pédagogique peut affecter l'artéfact « coévaluation » dans la mesure où certains élèves ne perçoivent pas l'utilité ou le sens de l'activité coévaluative en soi. L'analyse montre que ce phénomène peut avoir pour cause un malentendu concernant le fonctionnement de l'algorithme à l'origine du score de coévaluateur dans Moodle Atelier (*i.e.* conflit instrumental d'origine technique). Chez ces élèves, on peut également remarquer une perception conflictuelle de la coévaluation, opposant inter- et hétéro-évaluation. Plus précisément, l'analyse des données textuelles a montré que, pour de tels élèves coévaluateurs, la source de ce conflit instrumental se trouve dans la croyance que l'évaluation est un acte arbitraire : l'élève peine à faire le lien entre les attendus didactiques (compétences cibles) et les composantes docimologiques afférentes (métacompétences). *In fine*, on se trouve donc en présence d'un conflit instrumental dont l'origine est à la fois didactique, pédagogique et technique – et qui est susceptible d'entraîner un renoncement total ou partiel à la coévaluation, *i.e.* à l'artéfact pédagogique, ainsi qu'à l'artéfact didactique, auquel la coévaluation conditionne l'accès.

Pourtant, une combinatoire d'artéfacts régulateurs – *i.e.* synergie régulatrice – est susceptible de pallier, partiellement ou totalement, cette combinatoire conflictuelle. En effet, l'enseignant peut envisager des remédiations pour permettre aux élèves de mieux appréhender le lien entre attendus didactiques et composantes docimologiques afférentes, pour un effet rétroactif sur l'artéfact didactique. Selon les préconisations des enseignantes partenaires elles-mêmes, et à la lumière de nos résultats, une pratique régulière de la coévaluation permettrait d'opérationnaliser la dimension didactique conflictuelle précitée. Enfin, certaines fonctionnalités régulatrices de l'artéfact numérique rendent la coévaluation individualisée, non seulement possible, mais systématisable, en vue d'une « familiarisation à la tâche » (Reynaud, 2022, p. 100), avec là encore un effet rétroactif, sur l'artéfact pédagogique cette fois, ainsi que des bénéfices potentiels en termes d'apprentissages (artéfact didactique). On se trouve donc ici en présence d'une synergie régulatrice se rapportant à la fois aux artéfacts didactique, pédagogique et technique, et impliquant des relations de causalité réciproque, ainsi que des

effets rétroactifs. Le Tableau 32 rend compte de cette « dialectique instrumentale » que nous avons pu observer du point de vue des élèves.

Conflits instrumentaux observés	Artéfacts conflictuels	Artéfacts affectés	Artéfacts régulateurs
Problèmes de discernement docimologique	Artéfact pédagogique	Artéfact pédagogique	Artéfact technique (« Exemples de travaux »)
	Artéfact didactique		
Inhibitions socioaffectives	Artéfact pédagogique	Artéfact didactique	Artéfact technique (anonymisation)
Préconceptions négatives à l'égard de l'artéfact pédagogique et/ou perception arbitraire de l'évaluation	Artéfact didactique	Artéfact pédagogique	Artéfact didactique
	Artéfact pédagogique		Artéfact pédagogique
	Artéfact technique		Artéfact technique
Difficulté à formuler un écrit commentatif adapté	Artéfact didactique	Artéfact pédagogique (inter-évaluateur)	Artéfact didactique
		Artéfact didactique (pair évalué)	
Mésusages ou non-usage des interfaces	Artéfact technique	Artéfact pédagogique (inter-évaluateur)	Artéfact technique (familiarisation)
	Artéfact didactique	Artéfact didactique (pair évalué)	Artéfact didactique
Manque de prérequis didactiques	Artéfact didactique	Artéfact pédagogique	Artéfact technique

Tableau 32 – Principaux conflits instrumentaux identifiés et artéfacts afférents, du point de vue des élèves

Artéfact didactique	
Artéfact pédagogique	
Artéfact technique	

Tableau 33 – Légende

Nous avons pu observer d'autres processus de régulation instrumentale et avons procédé à une analyse croisée, pour déterminer quelles combinatoires conflictuelles ces différents facteurs de régulation sont susceptibles de réguler. A la faveur de cette analyse, on voit émerger d'autres combinatoires d'artéfacts régulateurs ou synergies régulatrices. On peut notamment citer la partie finale de la tâche coévaluative, qui consiste à produire un bref « écrit commentatif » (Brunel & Guérin-Callebout, 2018; Taous, 2020), *i.e.* une appréciation écrite

argumentée et constructive à destination des pairs coévalués. Nous avons vu que ce type de feedback entre pairs constitue un facteur potentiellement favorable à l'enseignement explicite et à l'autodirection des apprentissages (Gauthier *et al.*, 2007; Schunk, 1982), pouvant entraîner une augmentation du sentiment d'efficacité personnelle des élèves (Bandura, 2007; Berger & Büchel, 2012). Toutefois, un nombre significatif d'élèves coévaluateurs signalent la difficulté que représente pour eux cette partie de la tâche. Réciproquement, bon nombre de pairs coévalués estiment que ces écrits commentatifs ne sont pas toujours suffisamment explicites ou argumentés, probablement en raison d'une difficulté du pair coévaluateur à mobiliser les métacompétences requises, en termes de didactique de la langue. On se trouve potentiellement en présence d'une combinatoire conflictuelle d'origine didactique, affectant l'artéfact pédagogique pour le pair coévaluateur – mais aussi l'artéfact didactique pour le pair coévalué, chez qui cela peut entraîner un renoncement partiel à l'activité d'apprentissage. En l'occurrence, il s'agirait plutôt d'un renoncement partiel – et non d'un renoncement total –, dans la mesure où d'autres écrits commentatifs, rédigés par d'autres pairs coévaluateurs, sont *in fine* susceptibles de réguler ce conflit instrumental en fournissant un feedback constructif aux élèves concernés.

Ceci étant, plusieurs composantes artéfactuelles peuvent intervenir pour réguler ce conflit. Pour commencer, une remédiation didactique avec effet rétroactif sur l'artéfact didactique peut être mise en œuvre pour faciliter la mobilisation par les élèves des métacompétences requises pour rédiger l'écrit commentatif. De surcroît, la mise en œuvre d'une telle remédiation didactique peut faire l'objet d'une différenciation pédagogique synchrone, dans la mesure où elle peut prendre appui sur l'horizontalisation de la posture enseignante, induite par l'artéfact numérique. Enfin, un usage répété du dispositif expérimental de coévaluation permet aux élèves de se familiariser avec la tâche – ce qui peut créer un effet de synergie avec les deux facteurs de régulation instrumentale précités. En somme, une « dialectique instrumentale » est là encore à l'œuvre, avec des combinaisons d'artéfacts, liées par des relations de causalité réciproque et/ou des effets rétroactifs.

4.2 Des fonctionnalités régulatrices, apanage de l'artéfact numérique

Concernant la coévaluation dans son ensemble, et singulièrement pour les écrits commentatifs, les élèves participants ont insisté sur l'importance de l'anonymisation en double

aveugle – seul l’enseignant sait qui a écrit quoi, et pour qui. Une telle anonymisation permet d’éviter un conflit instrumental d’origine pédagogique en lien avec des biais d’ordre sociocognitif ou affectivo-motivationnel (Dochy *et al.*, 1999; Topping, 1998), gênant l’accès à l’artéfact didactique – et partant, aux métacompétences afférentes. Or l’anonymisation en double aveugle constitue une « fonctionnalité régulatrice », apanage de l’artéfact numérique, qui permet précisément de réguler ce conflit instrumental.

En effet, nous avons pu constater empiriquement l’intérêt pour les apprentissages de certaines fonctionnalités régulatrices spécifiques au numérique, *e.g.* l’anonymisation en double aveugle précédemment évoquée, mais aussi la scénarisation interactive, au sein de l’artéfact numérique lui-même, d’exercices d’appropriation des artéfacts pédagogique et didactique au moyen d’« Exemples de travaux », ou encore la génération d’un score de coévaluateur – fruit d’un algorithme opérationnalisant l’approche par compétence en comparant finement, critère par critère, les inter- et les hétéro-évaluations, pour fournir aux enseignants comme aux apprenants un feedback synthétique concernant les métacompétences de ces derniers.

A titre d’exemple, la fonctionnalité « Travaux exemplaires », qui est l’apanage de l’artéfact numérique et permet un entraînement à la coévaluation avec feedback immédiat – *i.e.* automatisé, en se basant sur une hétéro-évaluation de référence réalisée en amont par l’enseignant – peut contribuer à réguler plusieurs conflits instrumentaux précédemment évoqués, *e.g.* le manque de prérequis didactiques, la perception de l’évaluation comme acte arbitraire ou les inhibitions socioaffectives en lien avec la peur de « noter injustement » un pair.

4.3 Peut-on pratiquer la coévaluation individualisée sans artéfact numérique ?

Nous avons vu que certains processus de régulation étaient l’apanage du numérique, par le truchement de « fonctionnalités régulatrices ». A fortiori, pour un groupe classe de taille moyenne au collège – *i.e.* 30 élèves –, certaines des fonctionnalités régulatrices précitées peuvent constituer un prérequis pour pratiquer la coévaluation de manière individuelle. En effet, l’anonymisation manuelle d’une trentaine de productions écrites d’élèves constituerait une tâche chronophage, difficile voire impossible à généraliser dans le quotidien de la classe. De même, il nous semble que l’organisation concomitante d’un système manuscrit garantissant l’anonymat des évaluations entre pairs représenterait une gageure. Enfin, les opérations mathématiques permettant de comparer, compétence par compétence, les inter- et les hétéro-

évaluations pour aboutir à un score de coévaluateur nous paraissent impossibles, sinon à exécuter à une telle échelle, du moins à réaliser de manière suffisamment facile pour l'enseignant, de telle sorte qu'il soit enclin à pratiquer fréquemment cette coévaluation individuelle avec l'ensemble de sa classe. A cet égard, si la pratique manuscrite de la coévaluation par petits groupes permet de se passer des opérations d'anonymisation – dans la mesure où, eu égard à la production et à la coévaluation d'une production écrite, la dilution de la responsabilité individuelle au sein d'un collectif permet de limiter les biais sociocognitifs et les inhibitions socioaffectives –, il n'en demeure pas moins que la coévaluation par petits groupes ne permet pas aux élèves de découvrir les écrits individuels de leurs pairs, ni d'exercer leurs compétences métacognitives individuelles.

Dans l'ensemble, les résultats de la présente recherche permettent donc d'affirmer qu'un artéfact numérique comme Moodle Atelier – instrumentalisé au moyen de la catachrèse « coévaluation » détaillée *supra* – rend tout simplement *possible* la coévaluation individualisée en classe. Au-delà du fait de rendre cette coévaluation individualisée possible, nous avons vu que certains facteurs de régulation instrumentale peuvent contribuer à ce que la coévaluation individualisée soit mise en œuvre régulièrement, notamment en s'appuyant sur l'appétence des élèves interrogés, à la fois pour l'artéfact pédagogique (*e.g.* ludification de la métacognition, motivation intrinsèque liée à la dimension coopérative de l'activité) et pour l'artéfact numérique (*e.g.* facilité d'utilisation, clarté des consignes, « Exemples de travaux », anonymisation en double aveugle).

A cet égard, nous avons également pu constater qu'un cercle vertueux peut s'initier, dans la mesure où les élèves interrogés ont souligné l'importance de l'expérience acquise avec le dispositif – *i.e.* pour une « montée en métacompétences », confirmée par l'augmentation que nous avons pu établir concernant les scores de coévaluateurs des élèves participants. Or nous avons également pu montrer que les scores de scripteurs tendent à augmenter, *i.e.* les compétences des élèves participants en maîtrise de la langue et production écrite en français langue première tendent à s'améliorer. En résumé, la Figure 78 constitue une piste interprétative du processus qui semble à l'œuvre pour les élèves, lors des séances expérimentales avec le dispositif pédagogique numérique de coévaluation, relativement à sa genèse instrumentale – et au triple enchâssement des artéfacts didactique, pédagogique et technique qui la caractérise.

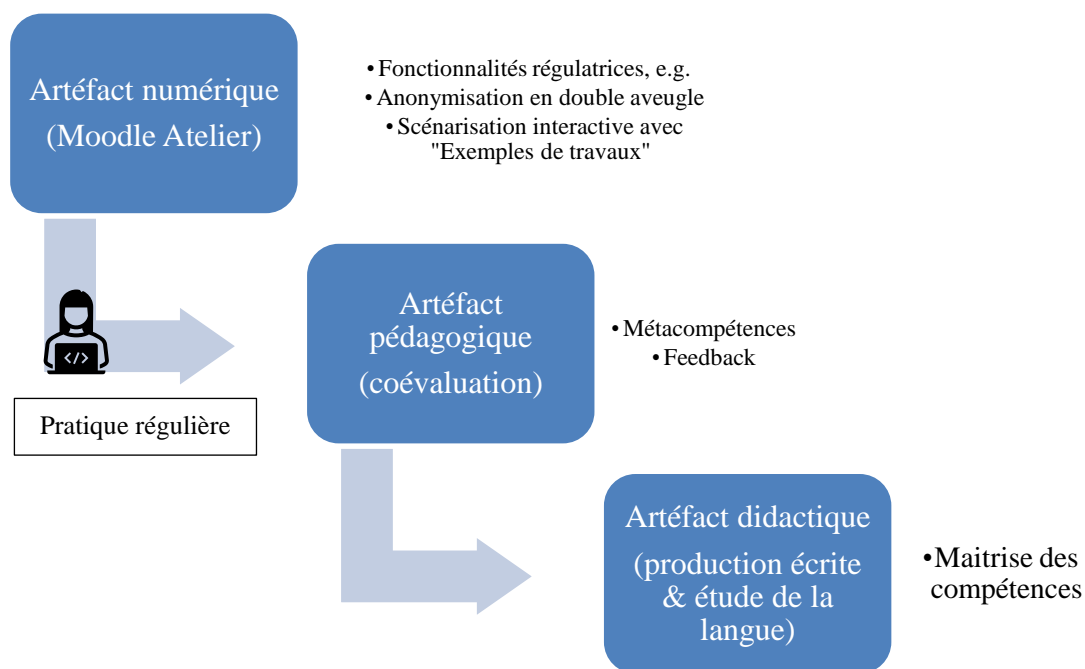


Figure 78 – Piste interprétative des interactions à l'œuvre dans la genèse instrumentale du dispositif

4.4 Degrés de conflictualité et effets sur les apprentissages

Si les processus de régulation qui sont à l'œuvre sont à même de compenser partiellement ou totalement un nombre significatif de conflits instrumentaux, il n'en demeure pas moins que ces processus de régulation sont à appréhender dans leur complexité, *i.e.* en tant que continuum. Cette notion de continuum appliquée aux processus de régulation instrumentale, s'inscrit dans le prolongement des travaux de (Marquet, 2003, 2005b, 2011; Marquet & Coulibaly, 2011) et de la hiérarchisation des conflits instrumentaux que nous avons pu proposer, selon une typologie distinguant entre les conflits instrumentaux entraînant un retard pour la situation d'enseignement/apprentissage, ceux qui entraînent un renoncement partiel et ceux qui entraînent un renoncement total à l'activité. Lorsque nous avons utilisé cette typologie pour affiner notre compréhension des conflits instrumentaux identifiés, nous avons constaté que les principaux conflits instrumentaux identifiés entraînent toujours des conséquences significatives sur les situations d'enseignement/apprentissage. Dans ce contexte, nos résultats permettent de supposer que tout processus d'ajustement instrumental est susceptible de réguler plus ou moins efficacement un conflit, pour le pallier partiellement ou totalement, selon un « continuum de la régulation instrumentale », à la faveur duquel un renoncement total peut devenir partiel, un renoncement partiel peut devenir un retard et un retard peut cesser. En vue de l'opérationnalisation d'un tel continuum de la régulation instrumentale, et eu égard aux combinatoires conflictuelles constatées dans le cadre de la

présente recherche, le Tableau 34 ci-dessous constitue une récapitulation des résultats présentés précédemment, en distinguant cette fois-ci entre artéfacts conflictuels et artéfacts affectés.

Conflits instrumentaux	Artéfacts conflictuels	Artéfacts affectés	Degrés de conflictualité		
			Retard	Renoncement partiel	Renoncement total
Problèmes de discernement docimologique	Artéfact pédagogique	Artéfact didactique	Retard	Renoncement partiel	
	Artéfact didactique				
Inhibitions socioaffectives	Artéfact pédagogique	Artéfact didactique	Retard		
L'évaluation perçue comme arbitraire	Artéfact didactique	Artéfact pédagogique	Retard	Renoncement partiel	
	Artéfact pédagogique				
Difficulté à formuler un écrit commentatif adapté	Artéfact didactique	Artéfact didactique	Retard	Renoncement partiel	
	Artéfact technique				
Contresens pédagogique à l'égard de la coévaluation	Artéfact didactique	Artéfact pédagogique (coévaluateur)		Renoncement partiel	Renoncement total
		Artéfact didactique (pair coévalué)			
Mésusages ou non-usage des interfaces	Artéfact technique	Artéfact pédagogique (coévaluateur)	Retard	Renoncement partiel	
	Artéfact didactique	Artéfact didactique (pair coévalué)			

Tableau 34 – Principaux conflits instrumentaux observés et degrés de conflictualité afférents (point de vue élèves)

Artéfact didactique		Retard	
Artéfact pédagogique		Renoncement partiel	
Artéfact technique		Renoncement total	

Tableau 35 – Légende

4.5 Retour réflexif sur le point de vue des enseignantes partenaires

Du point de vue des enseignantes partenaires, les conflits instrumentaux rencontrés se rapportent essentiellement aux artefacts didactique et technique du dispositif expérimental, ainsi qu'à une combinatoire pédagogique et technique. Rappelons que chacune des trois enseignantes partenaires a étroitement collaboré avec le chercheur pour co-construire le protocole semi-expérimental d'instrumentalisation de Moodle Atelier pour la coévaluation, le mettre en œuvre avec ses élèves respectifs et prendre part, pendant toute la phase expérimentale en année N – et même en année N+1 (enseignante B) –, à un processus itératif d'ajustement instrumental continu dudit protocole. A la faveur de la posture épistémologique qui caractérise la présente recherche, à savoir son inscription dans le paradigme de la recherche collaborative en sciences de l'éducation et de la formation, l'approche instrumentale en environnement informatisé pour l'apprentissage humain constitue un cadre de réflexion commun pour le chercheur-intervenant et les enseignantes partenaires. En d'autres termes, cette partie de notre cadre théorique a d'emblée été explicitement partagée avec les enseignantes partenaires. Il s'agissait en effet de les associer en toute conscience à la co-construction des savoirs, dans le cadre d'une recherche dont elles étaient pleinement actrices – que ce soit lors de la phase exploratoire, lors de la phase expérimentale ou même relativement à la valorisation des résultats de la recherche (*e.g.* publications scientifiques ou dans des revues d'interface).

Pour rappel, les conflits instrumentaux qui s'accompagnent de processus de régulation instrumentale s'avèrent plus nombreux que les conflits instrumentaux irrésolus. Relativement aux régulations instrumentales observées, le Tableau 36 propose une synthèse des résultats présentés précédemment, en précisant pour chaque régulation observée les artefacts conflictuels, les artefacts affectés et les artefacts régulateurs. On constate que, bien souvent, ces artefacts sont liés par des relations de causalité réciproque – avec, dans certains cas, des boucles de rétroaction régulatrices, *e.g.* pour « Rapport d'évaluation et contrôle didactique », « Ecrit court plutôt qu'écrit long » ou « Restriction du nombre de critères d'évaluation ». En d'autres termes, un artefact initialement conflictuel peut être amené à jouer un rôle régulateur.

Régulations instrumentales observées	Artéfacts conflictuels	Artéfacts affectés	Artéfacts régulateurs
Ecrit court plutôt qu'écrit long	Artéfact didactique	Artéfact pédagogique	Artéfact didactique
Restriction du nombre de critères d'évaluation	Artéfact didactique	Artéfact pédagogique	Artéfact didactique
Etayage métacognitif au moyen d'échelles de Likert et de descriptifs ad-hoc	Artéfact didactique	Artéfact pédagogique	Artéfact didactique
Concertations spontanées entre élèves : débats docimologiques	Artéfact didactique	Artéfact pédagogique	Artéfact numérique (hybridité)
« Rapport d'évaluation » et contrôle didactique	Artéfact numérique	Artéfact didactique	Artéfact numérique
Scores exprimés en pourcentages pour déjouer le « totem » de la note, <i>i.e.</i> l'occultation des compétences cibles	Artéfact pédagogique	Artéfact didactique	Artéfact numérique
Pratiquer régulièrement la coévaluation pour développer la métacognition et la coopération	Artéfact pédagogique	Artéfact didactique	Artéfact pédagogique
			Artéfact numérique
Objectivation accrue des pratiques évaluatives	Artéfact pédagogique	Artéfact didactique	Artéfact pédagogique
			Artéfact numérique
Augmentation de la motivation intrinsèque	Artéfact pédagogique	Artéfact didactique	Artéfact pédagogique
			Artéfact numérique

Tableau 36 – Principales régulations instrumentales observées du point de vue des enseignantes partenaires, avec artéfacts conflictuels, artéfacts affectés et artéfacts régulateurs afférents

Artéfact didactique	
Artéfact pédagogique	
Artéfact numérique	

Tableau 37 – Légende

Les conflits instrumentaux et processus de régulation afférents, que nous avons pu identifier avec les enseignantes partenaires, mettent en évidence une évolution de leurs schèmes d'enseignement. A cet égard, lorsqu'elle s'est réapproprié le dispositif expérimental en année N+1 avec une autre classe, nous avons vu que l'enseignante B a proposé une innovation particulièrement intéressante vis-à-vis du protocole co-construit pour l'année N : il ne s'agit plus alors, pour l'ensemble des compétences évaluées, d'une seule et même échelle de Likert à quatre points – allant de « Insuffisant » à « Très satisfaisant » –, mais de descriptifs ad hoc, adaptés à chaque critère, avec deux, trois ou quatre niveaux de maîtrise selon les cas. Cet étayage métacognitif paraît avoir aidé les élèves à mieux s'approprier les critères d'évaluation et à mieux évaluer leurs pairs. A cet égard, lors de la première réappropriation autonome du dispositif par l'enseignante B, les résultats quantitatifs relatifs à la coévaluation pour l'échantillon B' ont révélé un score moyen de 45% pour la coévaluation, ce qui peut paraître inférieur aux attentes compte tenu de l'étayage mis en place. Cependant ce résultat légèrement inférieur à la moyenne peut être interprété comme une réussite relative, si l'on tient compte du fait qu'il s'agit d'une classe en difficulté, située dans un collège REP. De surcroît, nous avons pu constater de très bonnes performances de la part de ces élèves en tant que scripteurs, avec un score moyen de 80%. On peut donc supposer que, de manière rétroactive, l'étayage didactique mis en place pour la coévaluation (artéfact pédagogique) a, *in fine*, un effet positif sur l'artéfact didactique (*i.e.* maîtrise des compétences cibles).

A cet égard, on peut rappeler une autre évolution notable des schèmes d'enseignement, qui consiste en un processus d'instrumentation. En effet, la coévaluation avec Moodle Atelier – *i.e.* une des nombreuses activités possibles dans Moodle – a exercé un pouvoir d'entraînement, qui a amené l'enseignante B à insérer dans le cours Moodle l'ensemble de sa séquence d'enseignement. Cette instrumentation, *i.e.* ce cas de figure dans lequel c'est l'artéfact numérique lui-même qui exerce une influence sur les schèmes d'utilisation des artéfacts didactique et pédagogique par l'utilisateur, fait écho à des pratiques similaires observées chez l'enseignante A – quoique dans son cas Moodle faisait déjà l'objet d'une utilisation avancée, notamment dans le cadre des actions de formation des enseignants qu'elle co-anime.

Pour finir, un conflit instrumental rencontré par l'enseignante B lors de sa première réappropriation autonome du dispositif en année N+1 s'avère instructif, singulièrement lorsqu'on distingue ce qui relève de l'artéfact technique et ce qui relève de l'artéfact pédagogique – en l'occurrence, le choix d'une modalité distancielle asynchrone pour pratiquer la coévaluation numérique. A cet égard, il est intéressant de considérer rétrospectivement les propos de l'enseignante B lors de l'entretien semidirigé réalisé à T4 : à l'époque, elle faisait justement part de sa volonté d'expérimenter le dispositif en distanciel asynchrone à l'avenir, tout en émettant des réserves, précisément concernant la difficulté qu'un nombre significatif d'élèves pourraient avoir à s'engager dans ce type de devoir à faire à la maison. Après sa première réappropriation autonome du dispositif en distanciel asynchrone en année N+1, ses observations confirment son intuition d'alors. En outre, elles corroborent l'analyse des données quantitatives que nous avons recueillies et montrent que, si l'artéfact numérique Moodle Atelier peut s'accommoder d'une grande variété de schèmes d'utilisation – de par les nombreuses options de paramétrage qu'il permet –, son paramétrage n'en demeure pas moins une opération complexe et délicate, qui peut engendrer de nombreuses erreurs et générer divers conflits instrumentaux. A fortiori, nous avons vu qu'en distanciel asynchrone et avec des apprenants en difficulté, ces conflits instrumentaux sont susceptibles d'entraîner des retards d'apprentissage, voire des renoncements partiels ou totaux. Ce résultat tend donc à confirmer que la modalité hybride que nous avons privilégiée était la plus adaptée pour des collégiens.

Ce qu'il faut retenir

En somme, du point de vue des élèves comme de celui des enseignants, la genèse instrumentale du dispositif expérimental de coévaluation numérique révèle un nombre significatif de conflits instrumentaux. Bien souvent, ces derniers ne fonctionnent pas de manière isolée mais entretiennent des relations complexes les uns avec les autres, *e.g.* causalité réciproque ou boucle de rétroaction régulatrice.

Toutefois, ces combinatoires conflictuelles sont le plus souvent régulées par des processus d'ajustement, mobilisant un ou plusieurs artéfacts, *i.e.* des synergies régulatrices. Dans le cas de l'artéfact numérique, des fonctionnalités régulatrices, qui en sont l'apanage, peuvent jouer ce rôle de régulation instrumentale. A plus forte raison, nous avons vu que certaines de ces fonctionnalités constituent *in fine* des prérequis pour pouvoir pratiquer la coévaluation *individualisée* avec une classe de collège de taille moyenne.

Les conflits instrumentaux observés impliquent des conséquences significatives pour les situations d'enseignement/apprentissage. Les processus de régulation instrumentale susceptibles de les résoudre partiellement ou totalement s'inscrivent toutefois dans un continuum : leur efficacité est conditionnée par le degré de conflictualité des conflits instrumentaux afférents.

Un retour réflexif sur le point de vue des enseignantes partenaires, relativement à la genèse instrumentale du dispositif, nous permet de mieux comprendre la dialectique instrumentale qui est à l'œuvre ici. Dans le cadre de cette dialectique instrumentale, conflits, régulations et composantes artéfactuelles entretiennent des relations de causalité complexes, avec parfois des artéfacts régulateurs produisant un effet rétroactif sur un artéfact conflictuel. A la faveur de cette analyse, nous avons vu pour finir que certains schèmes d'enseignement ont évolué au cours de la période expérimentale – et même au-delà, en année N+1. Certaines de ces évolutions se sont avérées fructueuses, d'autres ont montré par exemple qu'une modalité asynchrone et distancielle peut susciter de nouveaux conflits instrumentaux. Il en ressort que la modalité hybride, privilégiée pour le protocole de la présente recherche, constituait l'approche la plus adaptée pour des collégiens.

5 Coévaluation numérique et différenciation pédagogique

5.1 Artéfact numérique et différenciation pédagogique

Nos résultats montrent que la mise en œuvre d'ajustements pédagogiques ou didactiques est facilitée par le dispositif, dans la mesure où un certain nombre d'opérations sont automatisées par le truchement de l'artéfact numérique. En effet, bien que cela ait pu leur paraître parfois difficile et/ou chronophage, nous avons constaté que les enseignantes partenaires automatisent certaines des interactions qu'elles doivent habituellement gérer elles-mêmes, en programmant en amont, dans l'artéfact numérique, des éléments d'étayage didactique et de scénarisation pédagogique. A cet égard, on peut parler d'instrumentation dans la mesure où les schèmes d'enseignement sont guidés par l'architecture logicielle de Moodle Atelier, *i.e.* concrètement, les étapes de paramétrage d'un Atelier Moodle sont autant de propositions faites à l'enseignant pour scénariser sa situation d'enseignement/apprentissage.

Or l'analyse des données textuelles recueillies auprès des élèves montre que, dans l'ensemble, les enseignantes partenaires étaient davantage disponibles pour aider les élèves qui en avaient besoin. Cet aspect a été fortement souligné par les élèves interrogés, qui insistent sur l'importance pour eux de l'aide fournie par leur enseignante. Cette horizontalisation de la posture enseignante, instrumentée par le dispositif de coévaluation numérique, est donc propice à un accompagnement plus individualisé des élèves – *i.e.* à une plus grande différenciation pédagogique en classe.

En outre, nous avons vu, avec l'enseignante A notamment, que le dispositif pédagogique permet de mobiliser avec profit des fonctionnalités régulatrices, *e.g.* la mise en œuvre de scénarios interactifs par l'entremise d' « Exemples de travaux », qui constituent des exercices d'appropriation pour lesquels les élèves vont bénéficier d'un feedback personnalisé, automatisé et immédiat. Il y a différenciation pédagogique dans la mesure où les élèves les plus avancés peuvent par exemple décider de n'évaluer qu'un seul de ces « Exemples de travaux », tandis que les élèves qui en éprouvent le besoin, ou pour qui l'enseignant juge que cela est préférable, pourront évaluer la totalité. En outre, cette fonctionnalité peut être configurée de telle sorte que les entraînements sur « Exemples de travaux » soient facultatifs. La différenciation pédagogique tient également ici au fait que le feedback automatisé que va recevoir l'élève est précis et reflète fidèlement sa maîtrise des métacompétences cibles, dans la mesure où ce feedback est issu de ce même algorithme, par le truchement duquel l'artéfact numérique effectue une comparaison pair-à-pair, compétence par compétence, entre l'inter-évaluation de l'élève et l'hétéro-évaluation de l'enseignant.

Nous avons constaté par ailleurs que le dispositif expérimental de coévaluation numérique – et singulièrement, l'artéfact numérique – induisent une horizontalisation de la posture enseignante. Celle-ci repose notamment, au dire des enseignantes partenaires, sur une gestion de classe facilitée par le dispositif, notamment du fait de l'anonymat – qui limite les problèmes d'ordre socioaffectif –, mais aussi en raison de l'hybridité du dispositif. D'une part, en effet, le dispositif est à la fois présentiel et distanciel, car bien que tous les élèves soient rassemblés dans une seule et même salle informatique, ils coopèrent par ordinateur interposé. D'autre part, pendant les 55 minutes que dure la séance de coévaluation, chacun peut avancer à son rythme, ce qui favorise une certaine flexibilité, sans pour autant creuser les inégalités scolaires. En effet, les élèves les plus en difficulté peuvent avoir pour objectif d'atteindre un niveau de maîtrise satisfaisant eu égard aux compétences de base (*i.e.* le Socle commun de connaissances, de compétences et de culture). Pour leur part, les élèves les plus en réussite peuvent chercher à dépasser le niveau de maîtrise requis pour des compétences

d'approfondissement. En outre, comme le rappelle (Forget, 2017b, p. 30), « le développement de toute expertise suit une courbe logarithmique ». Ainsi, la marge de progression de ces élèves devient infime, passé un certain seuil ; toutefois les progrès réalisés n'en demeurent pas moins porteurs d'enrichissements significatifs en termes de degré de maîtrise. Enfin, entre l'élève en grande difficulté et l'élève en grande réussite, nos résultats montrent qu'il existe un continuum dans le cadre duquel chaque élève est susceptible d'avancer à son rythme et selon les perspectives d'apprentissage précitées, par l'entremise du dispositif expérimental de coévaluation numérique. De ce fait, le dispositif s'avère propice à la mise en œuvre de la différenciation pédagogique en classe.

En effet, les élèves les moins avancés peuvent prendre leur temps pour accomplir la tâche, d'autant plus que ladite tâche peut avoir été configurée en amont pour être « taillée » à la mesure de leurs compétences – *i.e.* conformément à leur zone de développement proche (Vygotski, 1985) –, notamment en adaptant le nombre de coévaluations à effectuer en fonction des compétences de l'élève, ou en attribuant de préférence des travaux plus faciles (ou plus difficiles) à coévaluer, en fonction de la qualité linguistique des écrits et des compétences des élèves coévaluateurs. La corrélation que nous avons identifiée entre les scores de scripteurs et les scores de coévaluateurs montre que cette approche peut constituer un point d'appui porteur en termes de différenciation pédagogique. Par ailleurs, nous avons vu que les élèves les plus avancés peuvent se voir attribuer des tâches d'approfondissement en dehors de l'activité de coévaluation stricto sensu.

De surcroît, nous avons vu que le dispositif expérimental de coévaluation numérique a occasionné la mise en œuvre d'étayages divers : étayages didactiques, pédagogiques ou spécifiquement liés à l'artéfact numérique. Des verbatim consultables en Annexes, dans le Tableau 64, montrent l'importance de ces différents types d'étayage pour les élèves participants. Ils sont accompagnés des scores moyens de scripteur et de coévaluateur de chaque élève, exprimés en pourcentage, pour l'ensemble de la période expérimentale (T1-T4 ou T1-T6 selon les échantillons). Les scores de scripteurs des élèves concernés s'avèrent légèrement supérieurs à la moyenne (54% en moyenne). Les scores de coévaluateurs sont quant à eux relativement élevés (67% en moyenne) et indiquent que la différenciation pédagogique induite par le dispositif a été porteuse en termes d'apprentissages, singulièrement en ce qui concerne les métacompétences cibles.

Plus globalement, cette tendance concernant l'amélioration des métacompétences des élèves à la faveur des étayages instrumentés par l'artéfact technique se trouve confirmée par l'analyse quantitative des données numériques afférentes. Pour l'ensemble des élèves

participants (super-échantillon SE, N = 61), on constate en effet que les scores de coévaluateurs tendent à augmenter (Tableau 38 et Figure 79). Seul le 3^e quartile fait exception, *i.e.* les 25% d'élèves qui ont obtenu les scores de coévaluateurs les plus élevés. En ce qui concerne le 1^{er} quartile en revanche – *i.e.* les 25% d'élèves ayant obtenu les scores de coévaluateurs les moins élevés –, l'augmentation est particulièrement significative. Ainsi, non seulement le score de coévaluateur de l'ensemble des élèves tend à augmenter, mais ce phénomène est particulièrement marqué pour les élèves les plus en difficulté.

Statistiques descriptives		
	T2	T4
N	61	61
Manquants	0	0
Moyenne	60.3	63.3
Médiane	63	64
Minimum	2	25
Maximum	96	97
25-ième percentile	48.0	53.0
50-ième percentile	63.0	64.0
75-ième percentile	76.0	75.0

Tableau 38 – Augmentation du score de coévaluateur pour SE (N = 61) entre T2 et T4

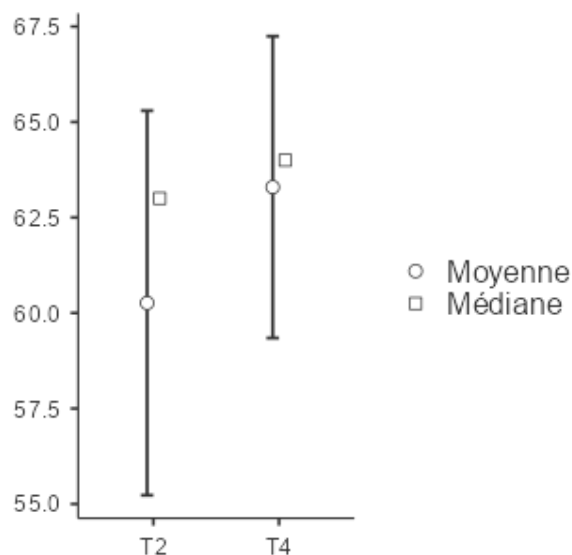


Figure 79 – Augmentation des score de coévaluateur pour SE (N = 61) entre T2 et T4

5.2 Différenciation pédagogique et métacognition

Pour rappel, selon (Grangeat, 1999, p. 126), l'articulation entre différenciation pédagogique et métacognition se traduit en classe par des dispositifs caractérisés par une

« conduite réfléchie des démarches d'apprentissage » requérant trois conditions. L'élève doit tout d'abord connaître les objectifs de l'enseignement (« enseignement explicite ») : or nous avons vu que la co-construction par les élèves et les enseignantes de la grille de coévaluation favorise cet enseignement explicite. Nous avons vu toutefois qu'un « contresens pédagogique » peut entraîner un renoncement partiel ou total en termes d'apprentissages. En effet, un conflit instrumental en lien avec l'artéfact pédagogique « coévaluation » peut survenir dans la mesure où certains élèves ne perçoivent pas l'utilité ou le sens de l'activité coévaluative en soi, ou entretiennent une vision conflictuelle entre inter- et hétéro-évaluation, qui repose sur la croyance que l'évaluation est un acte arbitraire. Cet élément constitue donc un point de vigilance.

Toutefois, les processus de réflexion docimologique – parfois spontanés –, que nous avons observés à la faveur du dispositif expérimental de coévaluation numérique, semblent de nature à réguler ce conflit instrumental. De même, nos résultats montrent que la pratique de la coévaluation numérique par l'entremise du dispositif expérimental (*i.e.* catachrèse pratiquée dans Moodle Atelier) contribue à une approche plus rationnelle vis-à-vis de l'acte évaluatif, pour les apprenants comme pour les enseignantes partenaires. En ce qui concerne les apprenants toutefois, cette rationalisation est conditionnée à une explicitation du fonctionnement de l'artéfact numérique et de son mode de calcul, singulièrement eu égard au score de coévaluateur – avec l'hétéro-évaluation de l'enseignant pour référence, et selon une comparaison compétence par compétence. Enfin, la fonctionnalité régulatrice « Exemples de travaux » semble à même de contribuer elle aussi à réguler le conflit instrumental précité, en familiarisant les apprenants avec cette tâche complexe qu'est la métacognition, avec un feedback immédiat, automatisé et personnalisé – mais sans la pression d'une « note qui compte dans la moyenne ».

5.3 Coévaluation numérique et inégalités socioscolaires : un bilan nuancé mais encourageant

Si l'on considère un des objectifs essentiels de la différenciation pédagogique, *i.e.* la réduction des inégalités socioscolaires – sans qu'il ne s'agisse bien évidemment de gommer toute différence (Crahay, 2013 ; Forget, 2017b) –, le bilan s'avère nuancé mais encourageant. D'une part, nous avons vu qu'un non-usage ou des mésusages des interfaces liées à l'artéfact numérique tendent à avoir lieu, notamment lorsque certains prérequis didactiques font défaut.

Ainsi, l'interprétation d'une suggestion de correction linguistique automatisée par le correcticiel, et le choix éclairé parmi plusieurs formes – aux graphies souvent proches –, peuvent être une source de confusion, singulièrement pour des élèves en difficulté.

On se souvient par ailleurs qu'avec l'enseignante B, l'expérimentation a été déployée auprès d'une classe de 3e, dans son collège situé en réseau d'éducation prioritaire (REP). Le numérique est envisagé par elle comme une ressource supplémentaire pour favoriser l'accès aux textes, aux mots, à l'écriture, pour des élèves souvent en (grande) difficulté – et pour essayer de réduire les inégalités socioscolaires. Le dispositif expérimental de coévaluation numérique constituait à cet égard, pour l'enseignante B, une proposition alternative, différente, potentiellement plus engageante pour des élèves peu motivés par les formes scolaires habituelles. Or les données recueillies auprès des élèves montrent que son intuition professionnelle s'avérait pertinente, notamment en raison de leur appétence pour l'artéfact pédagogique et sa dimension coopérative, de leur appétence pour l'artéfact numérique et en prenant appui sur la ludification de la métacognition, induite par le dispositif du point de vue des élèves – *i.e.* « c'est cool de faire le prof ».

Pour autant, pour les élèves les plus en difficulté – et pour qui le risque de décrochage scolaire est le plus grand –, des problématiques de motivation et d'autonomie viennent se surajouter. Y compris avec une activité comme celle-ci, qui a recours aux pédagogies du numérique pour proposer quelque chose de différent et de potentiellement engageant, nous avons vu que l'enseignante concernée déclarait n'obtenir que peu d'engagement, d'adhésion ou de prise au sérieux de l'activité de la part de ces élèves. Or les élèves qui étaient déjà en réussite au préalable profitent quant à eux du dispositif, et font des progrès ; ainsi, les écarts entre les élèves, en termes de maîtrise des compétences linguistiques, peuvent s'accroître. Nous avons vu toutefois que, relativement au degré de maîtrise d'une compétence, la courbe de progression est logarithmique et qu'au-delà d'un certain seuil, les progrès sont infimes tout en demeurant significatif en termes qualitatifs (*i.e.* enrichissement/approfondissement) ; de ce fait cette inquiétude ne semble pas fondée (Forget, 2017b).

Néanmoins, d'après les déclarations de l'enseignante B concernant le profil de quasi décrocheur scolaire de certains de ses élèves, et après examen des données quantitatives concernées – *i.e.* notes de scripteurs et de coévaluateurs – il s'avère que les élèves de l'échantillon B, dont la maîtrise de la langue est habituellement la plus fragile, et qui ont obtenu des notes particulièrement basses en tant que scripteurs, ont également obtenu des notes particulièrement basses en tant que coévaluateurs. En d'autres termes, si suivant l'avis de (Forget, 2017b) le présent dispositif n'est pas fondé à avoir contribué à accentuer les inégalités

socioscolaires, il semble qu'il n'ait pas non plus contribué à les réduire, *i.e.* en permettant aux élèves en grande difficulté d'acquérir une meilleure maîtrise des compétences cibles en français langue première.

On peut pourtant nuancer ce constat, notamment en prenant l'exemple d'un élève en grande difficulté de l'échantillon C. Sans revenir en détail sur l'exemple de cet élève, déjà détaillé *supra*, rappelons simplement que ce dernier – ci-après dénommé Etienne⁶⁸ pour les besoins de la discussion – faisait auparavant preuve d'une grande inhibition, avec en particulier un blocage face aux activités de production écrite. Or Etienne a manifesté un engagement fort dans les activités de production écrite et de coévaluation mises en œuvre dans le cadre du présent protocole. Singulièrement, on se souvient qu'il avait manifesté un remarquable « savoir-être » relatif à l'autorégulation des apprentissages, très peu présent chez lui auparavant d'après l'enseignante C.

Ainsi, l'artéfact numérique, qui a facilité la mise en œuvre de l'artéfact pédagogique « coévaluation », a également pu favoriser une maîtrise accrue des compétences liées à l'artéfact didactique « maîtrise de la langue et production d'écrit en français langue première ». En l'occurrence, les données quantitatives concernant Etienne montrent une augmentation significative et continue de ses performances de scripteur à T2, T4 et T6 (Figure 80).

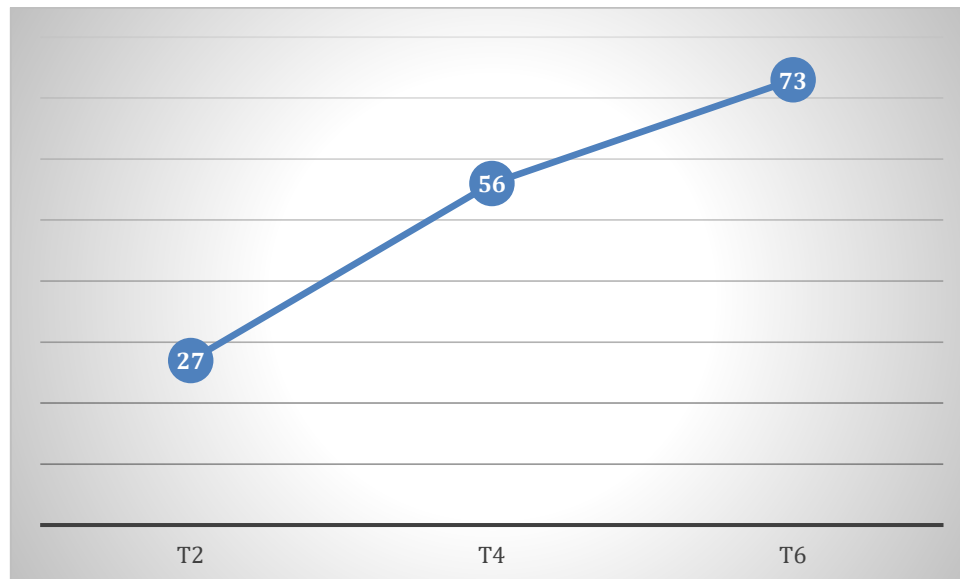


Figure 80 – Evolution du score de scripteur d'Etienne à T2, T4 et T6 (en %)

⁶⁸ Les prénoms des élèves ont été changés pour conserver leur anonymat, conformément au formulaire de consentement libre et éclairé signé par leurs parents et conformément au règlement général de protection des données (RGPD).

Les performances de coévaluateur d'Etienne n'ont pas suivi la même tendance. Après un score élevé à T2 (79 %) on constate une forte diminution à T4 (40%), avant une nouvelle augmentation à T6 (60%). La Figure 81 illustre cette évolution du score de coévaluateur d'Etienne.

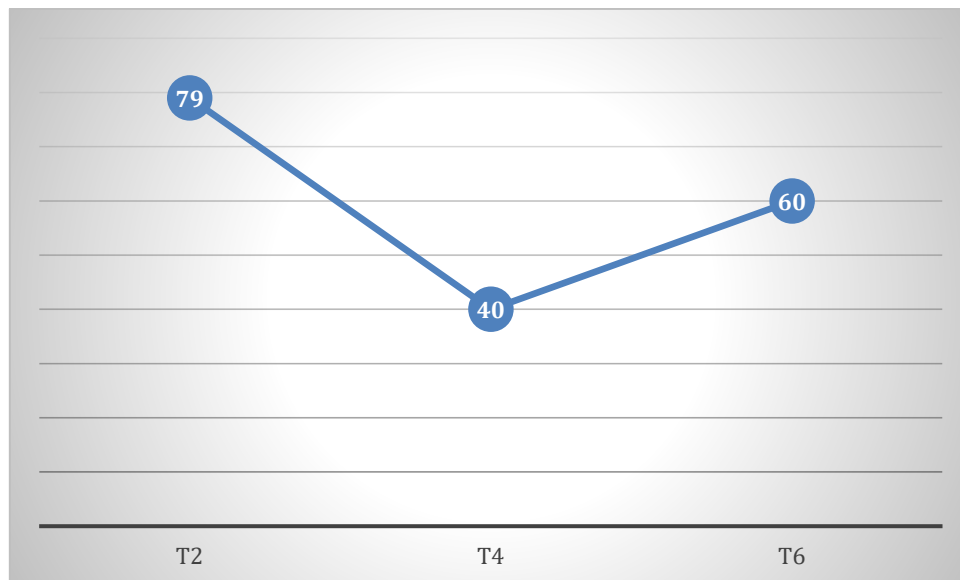


Figure 81 – Evolution du score d'évaluateur d'Etienne à T2, T4 et T6 (en %)

Toutefois, entre T1 et T6⁶⁹, on constate pour Etienne une augmentation puis une consolidation du sentiment d'efficacité personnelle en général (Figure 82). On constate la même tendance avec son sentiment d'efficacité personnelle pour l'autorégulation des apprentissages (Figure 83).

⁶⁹ En raison d'un problème technique survenu avec son poste informatique, nous ne disposons pas pour cet élève des données SEPE à T4, ni des données SEP(sr1) à T3 et T4.

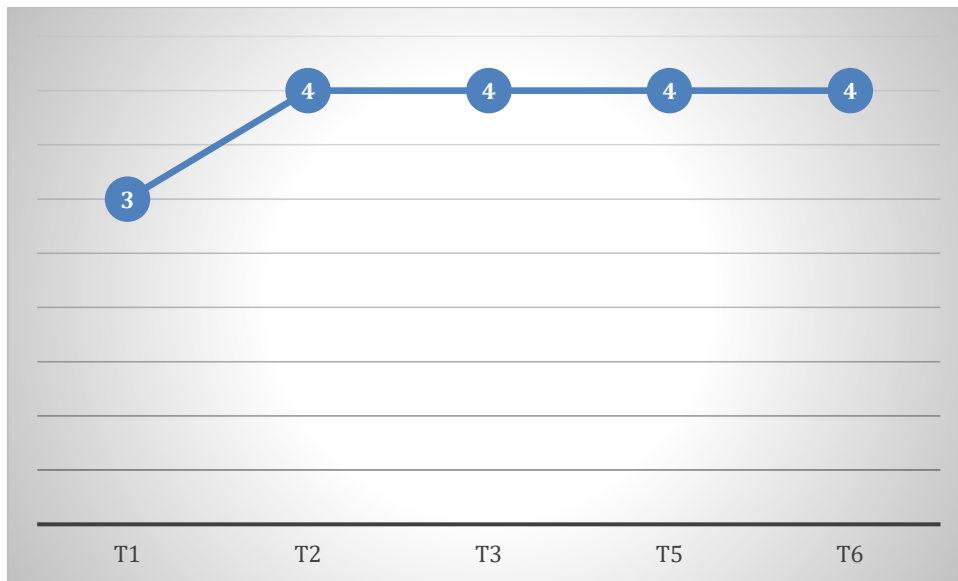


Figure 82 – Evolution du score individuel médian d'Etienne à l'échelle SEPE à T1, T2, T3, T5 et T6 (sur échelle de 1 à 5)

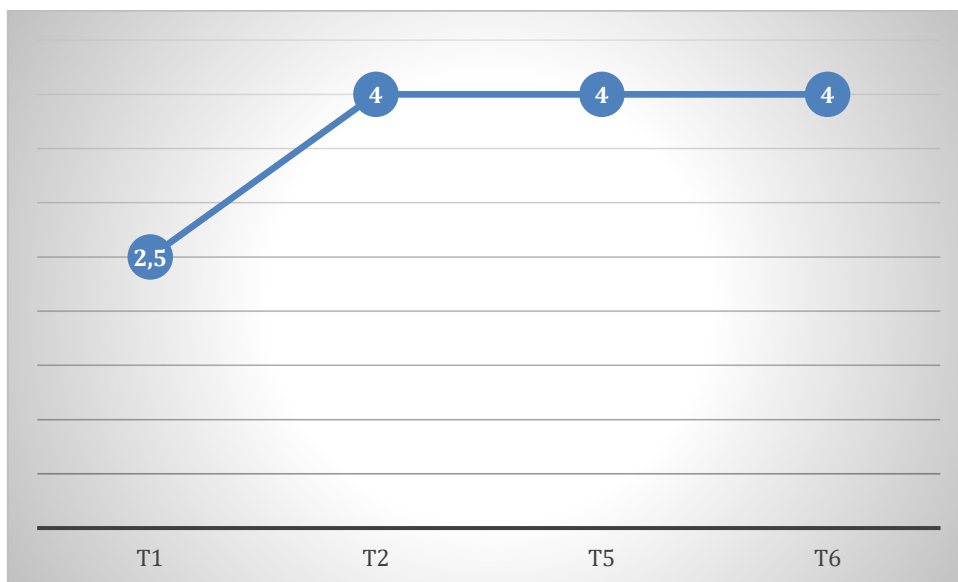


Figure 83 – Evolution du score individuel médian d'Etienne à l'échelle SEP(srl) à T1, T2, T5 et T6 (sur échelle de 1 à 5)

En résumé, cet apprenant en grande difficulté – souffrant d'inhibitions et de blocages pour la production d'écrit notamment, en lien avec une maîtrise jugée très fragile par l'enseignante C des compétences orthographiques et morphosyntaxiques mobilisées pour ce type de tâche – a accompli des progrès importants, à la fois en termes de compétences, de métacompétences et de sentiment d'efficacité personnel. Au dire de l'enseignante C, ces progrès sont à attribuer en grande partie à la manière dont Etienne s'est approprié le dispositif expérimental de coévaluation individualisée.

Bien qu'il s'agisse de cas particuliers, les exemples que nous venons d'analyser concernant les échantillons B et C nous paraissent revêtir une importance significative pour

comprendre la genèse instrumentale du dispositif expérimental de coévaluation et ses implications, à la fois en termes de différenciation pédagogique et de réduction des inégalités socioscolaires. Pour compléter ces analyses, on peut examiner les données numériques afférentes pour établir si, de manière générale, le dispositif a contribué à réduire les inégalités socioscolaires ou non. En d'autres termes, on peut s'intéresser aux indicateurs de tendance centrale et aux mesures de dispersion pour déterminer si, concrètement, le dispositif a permis de réduire les écarts entre les notes de scripteurs d'une part, et entre les notes de coévaluateurs d'autre part, pour l'ensemble des élèves (super-échantillon SE, N = 58) sur l'ensemble de la période expérimentale commune à tous les échantillons, *i.e.* entre T2 et T4.

Pour commencer, concernant les notes de scripteurs de SE (N = 58) à T2 et T4, on constate que, parallèlement à une augmentation significative observable via des indicateurs de tendance centrale comme la moyenne et la médiane (Figure 84), des mesures de dispersion comme l'étendue, l'écart-type et la variance diminuent substantiellement Tableau 39. En d'autres termes, les scores de scripteurs tendent à augmenter, tandis que les écarts inter-individuels semblent diminuer.

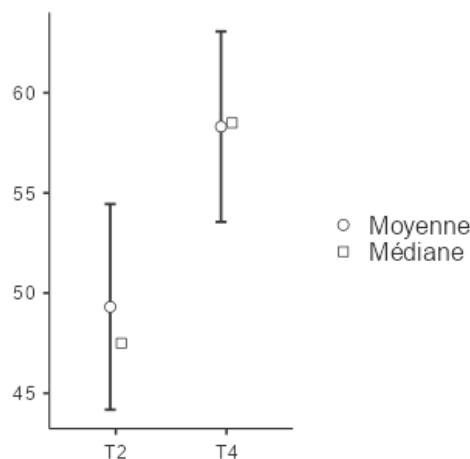


Figure 84 – Augmentation des scores de scripteurs de SE (N = 58) entre T2 et T4

Statistiques descriptives		
	T2	T4
N	58	58
Manquants	0	0
Moyenne	49.3	58.3
Médiane	47.5	58.5
Ecart-type	20.0	18.5
Variance	398	342
Ecart interquartile	26.3	28.0
Etendue	92.0	77.0

Statistiques descriptives		
	T2	T4
Minimum	8.00	17.0
Maximum	100	94.0

Tableau 39 – Notes de scripteurs de SE (N = 58) à T2 et T4 : statistiques descriptives

Toutefois ce constat est à nuancer. En effet l'écart interquartile (Figure 85) augmente significativement entre T2 et T4. Or une telle hausse peut être interprétée comme un accroissement léger mais non négligeable des écarts inter-individuels. On peut en conclure que l'amélioration globale des scores de scripteurs des élèves s'accompagne d'un léger accroissement des écarts inter-individuels, *i.e.* d'un léger accroissement des inégalités socioscolaires.

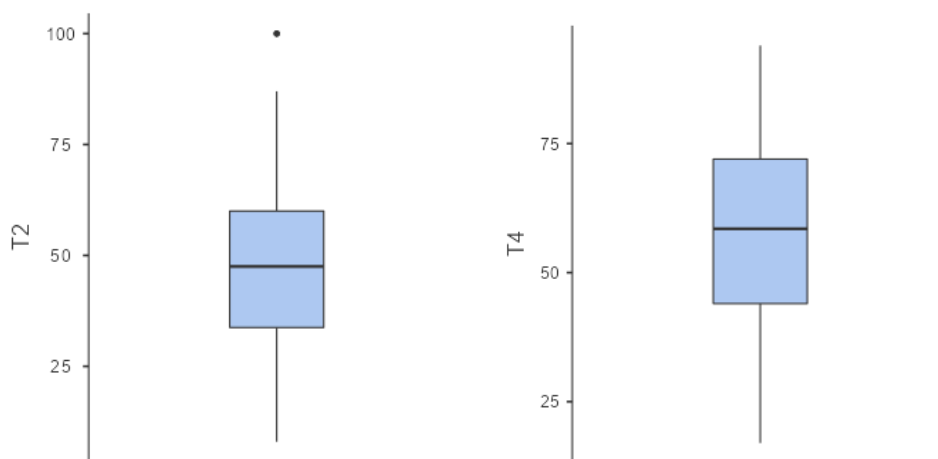


Figure 85 – Médiane et écart interquartile pour les notes de scripteurs de SE (N = 58) à T2 et T4

Lorsque l'on considère par ailleurs les notes de scripteurs de l'échantillon C (N = 22) à T2, T4 et T6 (Figure 86), on constate une augmentation significative et continue des notes moyennes de scripteurs.

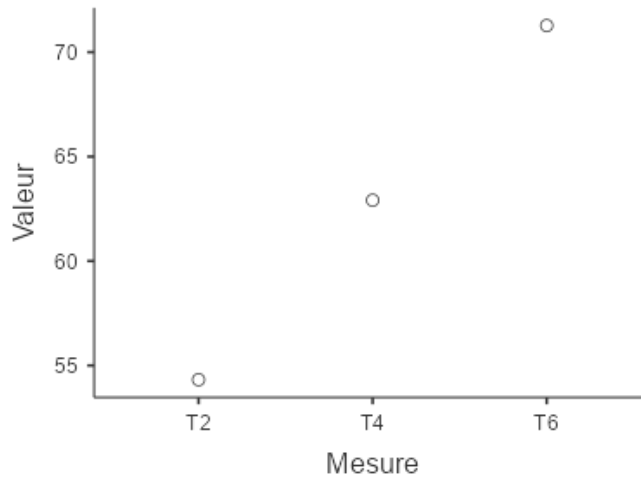


Figure 86 – Evolution du score de scripteur moyen de C (N = 22) à T2, T4 et T6

De manière concomitante (Tableau 40), on constate une hausse des indicateurs de tendance centrale et des indicateurs de dispersion – *e.g.* écart-type, variance et, dans une moindre mesure, écart interquartile (Figure 87).

Statistiques descriptives			
	T2	T4	T6
N	22	22	22
Manquants	90	90	90
Moyenne	54.3	62.9	71.3
Médiane	56.5	64.0	70.0
Ecart-type	17.3	17.5	19.7
Variance	300	305	388
Ecart interquartile	25.3	20.5	34.0
Etendue	60.0	77.0	60.0
Minimum	27.0	17.0	40.0
Maximum	87.0	94.0	100

Tableau 40 – Notes de scripteurs de C (N = 22) à T2 et T4 : statistiques descriptives

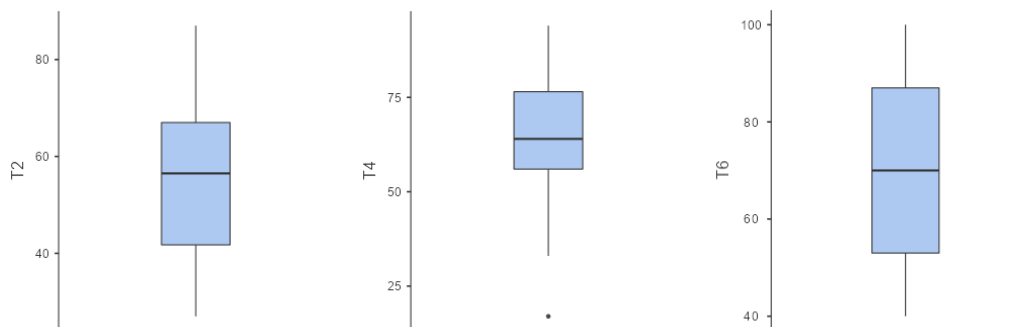


Figure 87 – Médiane et écart interquartile pour les notes de scripteurs de C (N = 22) à T2, T4 et T6

En d'autres termes, les scores de scripteurs tendent à augmenter, mais les écarts entre les scores individuels tendent à augmenter aussi. Là encore, l'amélioration des scores de scripteurs ne semble donc pas être assorti d'une diminution des écarts inter-individuels, *i.e.* d'une réduction des inégalités scolaires.

Après avoir examiné les notes de scripteurs des élèves, nous avons procédé à la même analyse concernant les notes de coévaluateurs. Pour SE (N = 61), entre T2 et T4, si l'on se réfère aux indicateurs de tendance centrale que sont la moyenne et la médiane, on constate là encore une augmentation globale des scores de coévaluateurs (Figure 88).

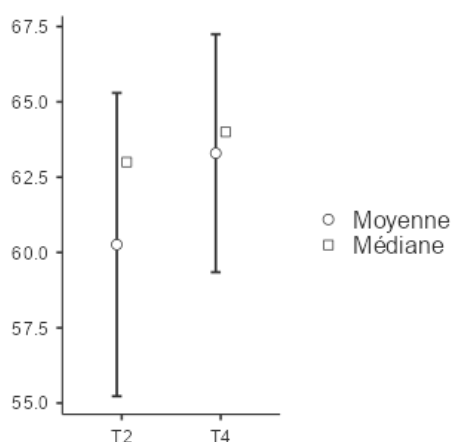


Figure 88 – Evolution du score de coévaluateur moyen de SE (N = 61) à T2 et T4

Cette fois cependant, on peut voir que les mesures de dispersion que sont l'étendue, l'écart-type, la variance ainsi que l'écart interquartile diminuent toutes substantiellement (Tableau 41). On peut donc affirmer que, non seulement le score de coévaluateur tend globalement à augmenter, mais les écarts inter-individuels tendent à se resserrer.

Statistiques descriptives		
	T2	T4
N	61	61
Manquants	0	0
Moyenne	60.3	63.3
Médiane	63	64
Ecart-type	20.1	15.7
Variance	402	248
Ecart interquartile	28.0	22.0
Etendue	94	72
Minimum	2	25
Maximum	96	97

Tableau 41 – Notes de coévaluateurs de SE (N = 61) à T2 et T4 : statistiques descriptives

Pour SE, pendant l'ensemble de la période expérimentale commune à tous les échantillons – *i.e.* de T1 à T4 –, on peut donc déduire que le dispositif expérimental a contribué à la fois à l'amélioration générale des performances de coévaluateurs et à la réduction des inégalités scolaires en la matière.

Pour finir, nous avons réalisé une analyse similaire avec les scores de coévaluateurs de l'échantillon C (N = 23) à T2, T4 et T6, afin de voir si la tendance observée pour SE sur T1-T4 se confirme pour C sur T1-T6. Les statistiques descriptives afférentes sont récapitulées dans le Tableau 42.

Statistiques descriptives			
	T2	T4	T6
N	23	23	23
Manquants	91	91	91
Moyenne	56.7	56.9	57.6
Médiane	58	58	60
Ecart-type	16.9	12.8	16.3
Variance	287	164	266
Ecart interquartile	23.0	21.0	14.5
Etendue	68	43	82
Minimum	16	35	2
Maximum	84	78	84

Tableau 42 – Notes de coévaluateurs de C (N = 23) à T2, T4 et T6 : statistiques descriptives

On constate que les indicateurs de tendance centrale augmentent continuellement pendant toute la période expérimentale « étendue » de l'échantillon C – *i.e.* T1-T6 –, à l'instar des scores moyens de coévaluateurs (Figure 89).

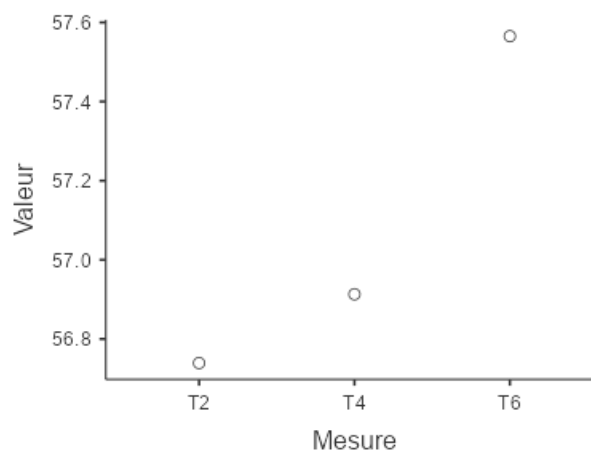


Figure 89 – Evolution des scores moyens de coévaluateurs de C (N = 22) à T2, T4 et T6

Par ailleurs, bien que des mesures de dispersion comme l'étendue, la variance ou l'écart-type diminuent à T4 avant de remonter à T6, l'écart interquartile diminue quant à lui substantiellement sans discontinuer pendant toute la période expérimentale (Figure 90), ce qui indique un resserrement des écarts entre les scores individuels de coévaluateurs des élèves de l'échantillon C, *i.e.* une réduction des inégalités scolaires en termes de métacompétences.

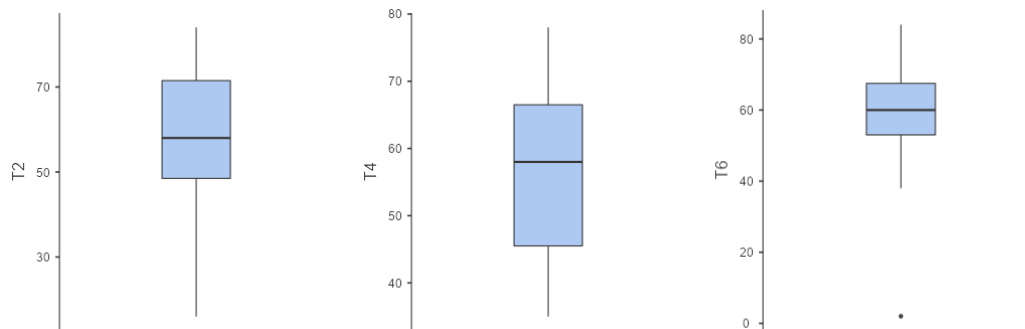


Figure 90 – Médiane et écart interquartile pour les notes de coévaluateurs de C (N = 23) à T2, T4 et T6

En somme, lorsque l'on considère l'ensemble des élèves participants et la totalité de la période expérimentale, ces analyses transversales qui convoquent, à divers degrés, l'intégralité de nos résultats et de notre cadre théorique, montrent un accroissement des écarts inter-individuels relativement aux scores de scripteurs. Il n'en va toutefois pas de même pour les scores de coévaluateurs : au contraire, on constate une amélioration globale et continue des métacompétences, assortie d'une réduction des inégalités socioscolaires en la matière. Compte tenu du fait que la tendance globalement haussière des deux types de scores en question – *i.e.* scripteurs et coévaluateurs –, est déjà établie par ailleurs au moyen d'analyses inférentielles, on peut supposer que, non seulement le présent dispositif permet une amélioration globale des performances de scripteurs et de coévaluateurs, mais il s'accompagne d'un resserrement des écarts inter-individuels, tout du moins eu égard au développement des métacompétences.

Ce qu'il faut retenir

Le dispositif expérimental de coévaluation numérique s'avère propice à une différenciation pédagogique accrue, soit grâce à l'horizontalisation de la posture enseignante qu'il induit structurellement, et qui rend l'enseignant davantage disponible pour aider ses élèves, soit par le truchement de fonctionnalités régulatrices qu'il instrumente, et qui favorisent la bonne compréhension des consignes et des attendus par les élèves. L'artéfact numérique offre également la possibilité d'ajuster la tâche à la zone proximale de développement de l'élève.

Nous avons vu par ailleurs que le dispositif expérimental satisfait aux critères définis par (Grangeat, 1999; Perrenoud, 2014) pour une articulation efficace entre métacognition et différenciation pédagogique. En outre, nous avons que les fonctionnalités régulatrices qui sont spécifiques à l'artéfact numérique permettent de mettre en œuvre divers types d'étayages aux effets porteurs, singulièrement en termes de métacompétences. Par ailleurs, l'analyse croisée des données numériques issues des différentes hypothèses opérationnelles testées a montré que le dispositif ne contribue pas à réduire les inégalités socioscolaires pour la maîtrise des compétences cibles en français langue première ; toutefois les perspectives semblent encourageantes dans la mesure où non seulement les compétences et les métacompétences des élèves tendent à s'améliorer, mais les écarts inter-individuels tendent à se resserrer eu égard aux métacompétences., *i.e.* le dispositif expérimental de coévaluation numérique contribue à réduire les inégalités socioscolaires dans ce domaine, tout en contribuant à une amélioration générale des compétences et des métacompétences.

6 La recherche collaborative comme retravail de l'empirie

La présente recherche a été mise en œuvre dans le cadre d'un processus collaboratif, tantôt synchrone, tantôt asynchrone, en distanciel et/ou en présentiel, entre les enseignantes partenaires et le chercheur. Cette modalité « co-élaborative » (Marcel, 2016b, p. 29) s'inscrit plus largement dans le paradigme de la recherche-intervention. Elle a favorisé un fructueux processus de concertation entre les enseignantes partenaires et le chercheur-intervenant, en associant pleinement ces dernières à la co-construction du protocole de recherche, « en s'appuyant sur leur expertise pour l'élaboration d'une réponse adaptée et acceptable » (Marcel, 2016b). Selon les dires des enseignantes partenaires elles-mêmes, cette recherche collaborative a permis de « proposer une tâche pertinente à la fois pour le projet de recherche et pour la progression de la classe » (Brulin *et al.*, 2023).

Ainsi, à la faveur d'un processus négocié, au carrefour entre les attendus scientifiques du chercheur-intervenant et les objectifs didactiques des enseignantes partenaires, le protocole initialement envisagé a évolué entre la phase exploratoire et le début de la phase expérimentale – et marginalement au cours de la phase expérimentale elle-même. Ce processus, intrinsèque à la démarche collaborative de recherche, consistait *in fine* en la co-construction par le chercheur et les enseignantes partenaires d'un « objet intermédiaire » qui « ne se réduit pas à l'intention

de son auteur » et dont « la matérialisation introduit quelque chose de nouveau, un glissement, qui n'est pas nécessairement voulu ni contrôlé [...]. L'objet résultant réalise autre chose que ce qui était voulu » (Vinck, 2009, pp. 56-57). Cet objet intermédiaire co-construit « interagit avec les acteurs en présence. Il supporte, par exemple, la confrontation de leurs points de vue en leur offrant des prises, en facilitant le surgissement de solutions et de rapprochements entre des aspects autrement dissociés » (Vinck, 2009, p. 59).

Par essence, le partenariat qui se constitue ainsi entre praticiens et chercheur-intervenant tend à exclure toute mise sous tutelle du dispositif de recherche et à le déborder. D'où la nécessité d'envisager des modalités de régulation, pour que ces débordements « ne remettent pas en cause les fondements mêmes du dispositif ». A cet égard, « la formalisation de la commande, au travers d'un contrat, peut fournir un certain nombre de repères » (Marcel, 2016, p. 30), dans la mesure où, « quand elle est possible, [cette dernière] revêt une fonction de cadre, qui à la fois protège et garantit la globalité fondatrice de la démarche » (Marcel, 2012, 2016, p. 28). Or la présente recherche ne s'inscrit pas dans le cadre d'une commande contractualisée (*e.g.* contrat doctoral, CIFRE⁷⁰). Toutefois, (Marcel, 2016, p. 26) signale que « la promotion d'un type particulier de thèse en mesure d'accueillir le format original de la démarche [collaborative] », est aujourd'hui une réalité, notamment au sein d'instances comme la 70^e section du CNU (Conseil national des universités), « en particulier pour la qualification des doctorants ».

A cet égard, ce travail de thèse nous permet de témoigner du fait que l'accompagnement des directeurs de thèse peut fournir un cadrage particulièrement efficace, y compris pour une recherche collaborative entreprise en dehors de toute commande contractualisée – comme c'est le cas ici. En particulier, à la faveur de la remise régulière de rapports, et par l'entremise des phases de restitution ou de communication en général (*e.g.* séminaires, colloques, articles), le suivi attentif et régulier des directeurs de thèse est susceptible de garantir la cohésion globale, et singulièrement la cohérence méthodologique, du projet de recherche, afin de limiter les risques de débordements évoqués par (Marcel, 2016b).

En outre, le fait que cette recherche semi-expérimentale soit menée en situation naturelle, en contexte scolaire, a impliqué une collaboration étroite avec les partenaires institutionnels concernés, *i.e.* les rectorats d'académie, les corps d'inspection de Lettres, les Cellules Académiques Recherche, Développement, Innovation et Expérimentation (CARDIE),

⁷⁰ CIFRE : Convention industrielle de formation par la recherche.

les délégations académique du numérique pour l'éducation (DANE et DRANE) de Paris et Montpellier, ainsi que les chefs des établissements concernés. Sans constituer une commande contractualisée stricto sensu, ces partenariats ont joué de facto un rôle de cadrage pour la présente recherche collaborative.

Par ailleurs, relativement à la dimension épistémique et à la circulation des savoirs, le chercheur-intervenant et les enseignantes partenaires ont pu expérimenter dans le cadre de la présente étude une dynamique interactionnelle féconde, bien que fondée sur des « savoirs spécifiques et différenciés entre les différents acteurs impliqués, savoirs qui ont leur légitimité propre, irréductibles l'un à l'autre » (Rhéaume, 2009, p. 204). Dans les faits, cette dialectique nous a permis d'expérimenter ce que (Marcel, 2016, p. 31) décrit comme un dépassement permettant « d'envisager la démarche de recherche-intervention, au-delà de ses objectifs propres, comme un espace de développement pour l'ensemble des acteurs concernés, chercheurs inclus ». Un développement sous-tendu par « les échanges, migrations et métissages de savoirs, de natures et de statuts différents, mais qui se voient mobilisés, déplacés ou confrontés au sein des diverses situations » (Marcel, 2016, *ibid.*). En guise d'exemple, on peut citer le témoignage de l'enseignante C à T6, en entretien semidirigé : à cette occasion elle a rendu compte de l'importance de l'approche par projet, qui a suscité une forte adhésion de la part des élèves. Ils avaient conscience de prendre part à un projet de recherche et de bénéficier à ce titre d'une attention accrue. En retour, les élèves de la classe C ont eu tendance à s'investir fortement dans les séances expérimentales, ce qui facilitait la dynamique de recherche, tout en constituant un biais potentiel (variable confondante). Lors d'échanges informels ultérieurs, par courriel – pour une approche des courriels comme écrits de travail, voir notamment (Torterat, 2015) – l'enseignante C est revenue sur ce sujet et a signalé la présence d'un potentiel effet de Hawthorne (Merrett, 2006; Roethlisberger & Dickson, 2003). Il nous semble que cet exemple manifeste la manière dont les savoirs des praticiens et du chercheur peuvent dialoguer de manière féconde.

Cette dialectique relative à la circulation de savoirs de types différents s'est avérée d'autant plus porteuse qu'elle s'inscrivait dans une dynamique communautaire, qui s'apparente aux concepts anglosaxons de *community-based research* (Jason & Glenwick, 2016), voire de *community-based participatory research* (Campano *et al.*, 2020; Zuber-Skerritt *et al.*, 2020), évoqués par (Pentecouteau, 2020, p. 11). En effet, bien que les enseignantes participantes enseignent toutes trois dans des établissements différents, dans des contextes socioculturels distincts – péri-urbain, rural, urbain en REP – et fortement éloignés géographiquement – puisque répartis entre Paris et Montpellier –, elles ont pu constituer avec le chercheur-

intervenant une communauté de recherche, certes essentiellement distancielle et asynchrone, mais efficacement opérationnalisée, par le truchement d'un projet qui s'est avéré fédérateur. Confrontées à des questionnements similaires, les enseignantes participantes ont ainsi pu échanger des informations et mutualiser des pratiques : des propositions d'activités de production écrite ou des suggestions docimologiques ont ainsi pu circuler au sein de cette communauté de recherche collaborative – y compris, au-delà de la phase expérimentale stricto sensu, *e.g.* via des réappropriations du dispositif en autonomie à N+1 ou par l'entremise d'une publication commune dans une revue d'interface (Brulin *et al.*, 2023). Cette publication, produite conjointement par les quatre acteurs, de manière distancielle et asynchrone, à la faveur d'un processus d'écriture collaborative particulièrement riche en soi, constitue d'ailleurs une manifestation concrète de ce que (Marcel, 2016a, p. 229) désigne comme « l'invention d'un langage capable non seulement de communiquer à l'intérieur d'une communauté plurilingue, mais permettant aussi de communiquer avec les langues d'origine de chaque partenaire » – *i.e.* un langage scientifique pour le chercheur-intervenant ou un langage professionnel pour les enseignantes partenaires.

Surtout, alors que (Marcel, 2016a, p. 229) souligne l'importance de la dimension spatiale, qui « dans sa dimension anthropologique, requiert également d'élaborer et d'habiter ensemble un tiers-espace de rencontre et d'action (voire de vie) de cette communauté chercheurs-acteurs », nous avons pu constater que, dans la réalité post-Covid qui est désormais la nôtre, une telle communauté peut néanmoins exister et se développer, au point de constituer ce que (Lescouarch, 2020, 2021) nomme un collectif apprenant.

Ce qu'il faut retenir

Selon une modalité hybride, la présente recherche-intervention s'inscrit dans le paradigme de la recherche collaborative en sciences de l'éducation et de la formation. De manière paritaire et horizontale, elle associe un chercheur-intervenant et trois enseignantes partenaires en s'appuyant sur l'expertise de chacun. A la faveur d'un processus de partage et de concertation, le protocole de recherche initialement prévu a ainsi évolué pour devenir un objet de médiation, co-construit, ayant vocation à interagir avec les acteurs en présence pour l'élaboration d'une réponse adaptée et acceptable.

L'inscription de cette démarche collaborative dans le cadre d'un projet de thèse permet de formaliser la commande, de la contractualiser et de fournir des repères pour en garantir la cohérence globale et éviter tous risques de débordements. Dans cette perspective, le suivi

des directeurs de thèse joue un rôle clé. De fait, outre l'université, pour une telle recherche collaborative en milieu scolaire, le cadrage est également institutionnel : il implique une étroite collaboration avec des partenaires tels que les rectorats d'académie, les corps d'inspection, les Cellules Académiques Recherche, Développement, Innovation et Expérimentation (CARDIE), les délégations (de région) académique du numérique pour l'éducation (DANE/DRANE), ainsi que les chefs d'établissements.

Le chercheur-intervenant et les enseignantes partenaires ont par ailleurs pu expérimenter ici une dynamique interactionnelle féconde eu égard à la circulation et au métissage de savoirs pourtant spécifiques et différenciés, parfois irréductibles mais tous également légitimes. *In fine*, les objectifs propres du chercheur-intervenant d'une part, et de chaque enseignante partenaire d'autre part, se conjuguent à travers un processus de dépassement, débouchant sur un espace de développement pour l'ensemble des acteurs concernés.

Ces apports mutuels sont fondés en premier lieu sur la manière dont les savoirs des acteurs et du chercheur-intervenant peuvent dialoguer de manière féconde. Ils relèvent d'une dialectique d'autant plus porteuse qu'elle s'inscrit dans le cadre d'une dynamique communautaire. A cet égard, les élèves et les enseignantes participantes ont pu constituer avec le chercheur-intervenant et les partenaires institutionnels une authentique communauté de recherche, au sens de la *community-based research* anglosaxonne. Ce fait semble d'autant plus encourageant qu'il implique trois professeures enseignant dans des établissements différents, dans des contextes socioculturels distincts – péri-urbain, rural, urbain en REP – et fortement éloignés géographiquement. La confrontation à des questionnements similaires a suscité des échanges d'information et une mutualisation des pratiques. De surcroît, un article a été coécrit par les enseignantes partenaires et le chercheur. Publié dans une revue d'interface, il est le produit d'une collaboration active et le témoignage d'un métissage fructueux entre des savoirs et des langages distincts.

Enfin, dans la réalité post-Covid qui est désormais la nôtre, la fécondité de la communauté de recherche engendrée par la présente étude témoigne du potentiel du paradigme de la recherche collaborative en sciences de l'éducation et de la formation, y compris selon une modalité hybride, impliquant des échanges distanciels/présentiels et synchrones/asynchrones.

7 Limites de cette étude

Avant de conclure, il s'agit d'évoquer dans cette section les limites que nous avons pu identifier pour ce travail de thèse.

7.1.1 Le correcticiel comme impensé de la présente recherche

On peut considérer que l'usage du correcticiel constitue une sorte d'impensé de la présente recherche, dans la mesure où l'équipement des traitements de texte au moyen d'un correcticiel – *e.g.* le logiciel libre Grammalecte – n'a pas été pris en compte. Or au-delà de l'équipement proprement dit, un correcticiel implique une littératie numérique ad hoc – *i.e.* une capacité à en faire usage à bon escient –, qui ne va pas de soi mais qui s'enseigne/s'apprend en classe, ainsi que le montrent les travaux de (Beer-Toker *et al.*, 2009; Delbrassine, 2018; Mangenot, 2012). Pour que l'usage du correcticiel puisse effectivement devenir un point d'appui dans le cadre de cette activité de production – puis de coévaluation – d'écrit instrumentée par le numérique, il conviendrait donc de développer davantage cet aspect dans de futurs protocoles de recherche, singulièrement pour investiguer la dimension didactique du dispositif pédagonumérique expérimental.

7.1.2 Variabilité du protocole

La variabilité du protocole semi-expérimental peut être considérée comme une limite de la présente recherche. Elle est liée d'une part à la situation naturelle, qui impliquait des contraintes d'ordre institutionnel et/ou organisationnelles pour les enseignantes partenaires – *e.g.* une progression didactique à maintenir, ou l'imminence des épreuves du diplôme national du brevet (DNB). Par ailleurs, si le protocole a évolué, c'est essentiellement pendant la phase exploratoire. Les aménagements qui ont eu lieu pendant la phase expérimentale, correspondaient à une nécessaire appropriation du dispositif par chaque enseignante partenaire. Ces variations étaient marginales dans la mesure où globalement c'est le même protocole qui a été mis en œuvre lors de toutes les itérations expérimentales, avec tous les échantillons – en particulier concernant les fondements du protocole, eu égard aux artéfacts didactique, pédagogique et technique du dispositif expérimental de coévaluation numérique. A fortiori, ces variations marginales étaient consubstantielles à cette recherche dans la mesure même où elle est collaborative et implique la co-construction d'un « objet intermédiaire » au sens de (Vinck,

2009), cité par (Marcel, 2016b) : en l'occurrence, cet objet intermédiaire est ici le protocole de recherche co-construit par les enseignantes partenaires et le chercheur-intervenant.

7.1.3 Longueur du questionnaire SEPE

Le questionnaire SEPE s'est avéré fiable, en raison de l'échelle de Likert à cinq points qui y est utilisée, des descriptifs afférents et du nombre élevé d'items ($N = 37$), qui permettent de mesurer, outre le sentiment d'efficacité personnelle des élèves participants, plusieurs sous-indicateurs qui se sont révélés très utiles pour notre étude. De plus, ce questionnaire était conçu spécialement pour le public cible. Pour autant, un nombre significatif d'élèves ont manifesté leur ennui, leur agacement ou leur incompréhension face à un questionnaire en ligne certes présenté via une interface ludique et ergonomique, mais déjà long à compléter en soi (5-10 minutes), et réitéré qui plus est avant et après chaque itération du protocole expérimental. Pour rappel, l'objectif de cette approche pré/post systématique était de compenser le délai potentiellement conséquent entre deux itérations du protocole de recherche – *e.g.* entre T1/T2 et T2/T3. Or il s'est avéré que les diverses contraintes précédemment évoquées, en lien avec la situation naturelle en contexte scolaire, occasionnaient parfois un délai de plusieurs mois entre deux itérations du protocole. En somme, bien que méthodologiquement ce procédé se soit révélé utile pour recueillir des données fiables et pertinentes pour tester nos hypothèses, il semble qu'une échelle moins longue eût été préférable.

7.1.4 Nombre réduit d'échantillons

Nous avons été confronté à des problématiques de recrutement, liées notamment au fait qu'il s'agissait d'une recherche collaborative. Il s'est en effet avéré difficile de trouver des enseignants volontaires, professeurs de français en cycle 4, disposé à s'investir à long terme dans un projet riche, mais susceptible d'être chronophage, puisque les enseignants étaient appelés à être, non pas simplement des participants, mais des partenaires. En outre, la pandémie de Covid-19 et les incertitudes afférentes ont rendu ces enrôlements encore plus compliqués en 2020, ce qui explique que nous n'ayons pu enrôler que trois enseignantes partenaires, dont l'une a fait participer deux de ses classes, soit un total de quatre classes/échantillons⁷¹.

⁷¹ La classe B' en année N+1 n'est pas comptabilisée ici dans la mesure où l'expérimentation s'effectuait alors hors protocole.

7.1.5 Nombre restreint d'itérations du protocole expérimental

A la fois en raison de la situation naturelle en contexte scolaire et des contraintes organisationnelles qui y étaient liées, et en raison de la pandémie de Covid-19, l'organisation des séances expérimentales pendant l'année scolaire 2020-2021 fut plus compliquée que prévu. Concrètement, les fenêtres d'opportunité furent moins nombreuses et soumises aux aléas des confinements successifs, ce qui explique le nombre restreint d'itérations du protocole expérimental : deux itérations avec chacun des échantillons A, A' et B, et trois itérations avec l'échantillon C, soit un total de neuf itérations, auxquels s'ajoutent les deux expérimentations autonomes menées par l'enseignante B en année N+1 avec l'échantillon B'. En outre, certaines des données textuelles n'ont pu être recueillies qu'auprès de l'échantillon C, avec qui trois itérations avaient été mises en œuvre – versus deux itérations avec tous les autres échantillons ; les résultats ainsi obtenus, et sur lesquels nous nous sommes appuyés pour une partie de nos conclusions correspondent donc à une réalité locale, au sein d'un échantillon déjà restreint (trois enseignantes, quatre classes, 80 élèves participants au total).

7.1.6 Taille restreinte des échantillons après appariement

Les tests non-paramétriques (analyses inférentielles) que nous avons été amené à pratiquer impliquaient un appariement des échantillons entre les différents temps de mesure, pour pouvoir être comparés. Pour être pris en compte statistiquement, l'élève devait avoir été présent à tous les temps de mesure concernés : *i.e.* à T1-T2, T3-T4 ou T5-T6 pour les analyses dites transversales, et de T1 à T4 ou de T1 à T6 pour les analyses longitudinales. Dans cette situation naturelle en contexte scolaire, l'absentéisme, même « perlé » – *i.e.* intermittent – a donc des conséquences statistiques significatives et a contribué à réduire la taille des échantillons appariés.

7.1.7 Utilisation de l'hétéro-évaluation des enseignants et biais potentiels

Pour les performances de scripteurs en particulier, nous avons choisi de nous appuyer sur les hétéro-évaluations des enseignants, malgré tous les biais sociocognitifs que cela peut impliquer (El Youssifi *et al.*, 2022; Mottier Lopez & Allal, 2008). En outre, l'artéfact numérique permettait d'anonymiser les productions écrites des élèves de telle sorte que les enseignants n'auraient pas su qui ils évaluaient. Aucun des enseignants partenaires n'a souhaité fonctionner de cette manière anonymisée, bien qu'elle eût été mise en œuvre pour leurs élèves par ailleurs. Cette approche de l'hétéro-évaluation correspondait à la pratique

habituelle en situation naturelle et nous avons donc jugé préférable de ne pas en altérer le fonctionnement. Surtout, s'il y a des biais sociocognitifs dans l'hétéro-évaluation des enseignants, on peut considérer que cela ne posait pas de problème méthodologique pour cette recherche dans la mesure où, *in fine*, ce sont les mêmes enseignants qui ont évalué les productions écrites des mêmes élèves tout au long de la phase expérimentale, avec potentiellement les mêmes biais sociocognitifs, ce qui en fait un instrument de mesure partiellement subjectif, mais fiable dans la mesure où cette « subjectivité contrôlée » se répétait d'itération en itération. Or nos analyses sont essentiellement des analyses *intragroupes*, entre temps de mesure, pour observer l'évolution des indicateurs au fil des itérations : nous n'avons délibérément pas effectué d'analyses *intergroupes*, précisément pour prendre en compte une variabilité intergroupe nécessairement substantielle dans le contexte de ce dispositif semi-expérimental en milieu naturel/contexte scolaire – singulièrement en ce qui concerne les biais sociocognitifs liés à l'hétéro-évaluation enseignante.

7.1.8 Effet de Hawthorne

En particulier avec l'échantillon C et grâce aux observations de l'enseignante C, nous avons pu identifier un surcroît de motivation de la part des élèves participants, dans la mesure où ils étaient conscients de participer à une expérimentation et appréciaient l'attention accrue que cela impliquait pour eux, notamment en termes de valorisation de leur travail. Ce phénomène, qui constituait à la fois un adjuvant pour la présente recherche, et une potentielle variable confondante, a été amplement documenté (Merrett, 2006; Roethlisberger & Dickson, 2003) dans le cadre de ce qu'il convient d'appeler l'effet de Hawthorne.

Ainsi, loin de prétendre à une quelconque portée générale, en premier lieu en raison de la taille limitée de la population participante et du nombre restreint d'itérations réalisées, il nous semble que les résultats de la présente recherche ne peuvent de surcroît pas faire l'économie d'une prudente relativisation, eu égard à l'effet de Hawthorne précité.

Pour limiter ce biais dans de futures recherches, on pourrait envisager que les enseignants partenaires soient formés en amont, afin de mettre en œuvre le protocole semi-expérimental en autonomie lors de la séance en classe – *i.e.* en l'absence du chercheur-intervenant –, et surtout sans que les élèves aient conscience de prendre part à un projet de recherche.

7.1.9 Difficultés de paramétrage et perspectives ergonomiques en lien avec l'artéfact numérique

Pour finir, il faut souligner que la catachrèse qui est au cœur de la présente recherche, et au moyen de laquelle nous avons pu instrumentaliser Moodle Atelier – *i.e.* un artéfact numérique initialement prévu pour un usage en inter-évaluation –, afin de pratiquer la coévaluation, consistait en une séquence de paramétrage avancé comportant de nombreuses étapes, souvent difficiles d'accès. Cette catachrèse est donc difficilement reproductible, du moins sans accompagnement. De ce fait, le dispositif expérimental de coévaluation numérique qui fait l'objet de la présente recherche ne semble pas susceptible de se diffuser largement, en tout cas de manière spontanée ou aisée. Certes, nous avons pu établir que le dispositif était porteur en termes d'apprentissages, à divers égards ; nous avons montré de surcroît que, sans un artéfact numérique comme Moodle Atelier, il n'est tout simplement pas envisageable de mettre en œuvre la coévaluation individualisée à l'échelle d'une classe standard de 30 élèves – et qu'à fortiori il serait impossible de pratiquer cette modalité évaluative de manière régulière sans recourir à un tel artéfact technique.

Les difficultés précitées peuvent cependant être nuancées. En effet, le dispositif expérimental de coévaluation numérique constitue dans une certaine mesure le cœur du protocole de recherche co-construit – *i.e.* le cœur de cet « objet intermédiaire » dont parle (Vinck, 2009) : initié par le chercheur-intervenant, il s'est enrichi du savoir professionnel et expérientiel des enseignantes partenaires. Nous avons d'ailleurs vu que les élèves participants – eux aussi acteurs à part entière de cette recherche collaborative – ont suggéré une catachrèse supplémentaire, susceptible d'enrichir la première pour mettre en œuvre une boucle formative de rétroaction anonymisée. En outre, l'enseignante B s'est réapproprié le dispositif en année N+1, en l'adaptant à ses besoins didactiques et à ses scénarios pédagogiques et en y implémentant pour ce faire, sans l'aide du chercheur-intervenant et avec succès, des schèmes d'utilisation nouveaux. A la faveur de la présente recherche et des médiations entre les différents acteurs du projet, cet « objet intermédiaire » a donc fait l'objet d'une adoption réelle, au sens de l'adoption de l'innovation telle qu'elle est théorisée par (Hall & Hord, 2006) notamment.

Dans ce contexte, *i.e.* une fois franchis un certain nombre d'obstacles pour s'approprier un outil nouveau, à la faveur des processus conflictuels et régulateurs qui surviennent à l'usage, relativement à la genèse instrumentale de ce dispositif en EIAH, il convient de signaler que l'artéfact numérique Moodle Atelier – qui s'inscrit plus généralement dans l'artéfact numérique

de la plateforme Moodle – offre des possibilités intéressantes pour un enseignant en termes de répliquabilité : tout atelier créé est en effet duplicable à l’infini dans Moodle et exportable pour d’autres classes et d’une année sur l’autre par exemple.

En dehors de cela, et sans recourir à une formation ad hoc, *e.g.* dans le cadre de la formation continue des personnels enseignants, il paraît cependant difficile, compte tenu de la complexité du paramétrage de la catachrèse, évoquée précédemment, de diffuser l’usage du dispositif expérimental de coévaluation numérique auprès d’autres enseignants. A titre d’exemple, à la demande de la cheffe d’établissement du collège A, nous avons animé en fin d’année un atelier « main à la pâte » qui a permis à une dizaine de collègues volontaires de tester le dispositif en jouant le rôle de l’élève. Le focus groupe qui a suivi a confirmé que la complexité du paramétrage de la catachrèse « coévaluation » était de nature à susciter pour les enseignants potentiels utilisateurs un conflit instrumental entraînant un renoncement total à l’usage du dispositif.

De fait, si Moodle présente l’avantage d’être un logiciel libre (non-propritaire), déjà installé dans tous les collèges publics de France, il n’en demeure pas moins qu’il est encore très peu connu et utilisé par les collègues enseignants. Je me souviens personnellement d’une formation portant sur Moodle, que j’avais suivie en 2010 dans le cadre de la formation continue des personnels enseignants lorsque j’enseignais en Australie. Presque quinze ans plus tard, force est de constater que cet outil, pourtant très utilisé dans l’enseignement supérieur français et pour la formation continue des enseignants, n’est que très peu mobilisé par les collègues de collège pour leurs élèves.

En somme, Moodle est une plateforme logicielle libre dotée d’une grande flexibilité en termes de potentialités d’usage. Cette plateforme est forte d’une communauté d’usagers et de développeurs bénévoles très dynamique – comme en témoigne notamment le succès des rassemblements « Moodle Moot », carrefours de la recherche et de l’innovation en e-learning avec Moodle. Toutefois, bien que cette flexibilité en termes d’usages – et à fortiori en termes de détournements d’usages –, confère à l’artéfact numérique Moodle Atelier un potentiel pédagogique virtuellement infini, cette plasticité va de pair avec une certaine exigence, compte tenu de l’effort requis pour en pénétrer les rouages et en comprendre les mécanismes de paramétrage.

Ce qu'il faut retenir

Parmi les limites de la présente étude, on peut évoquer en premier lieu l'absence de prise en compte du correcticiel, qui constituait pourtant un élément clé dans l'usage du traitement de texte, avec des répercussions significatives en production écrite mais aussi en réception/coévaluation. Par ailleurs la variabilité du présent protocole peut être considérée comme une source de biais pour les résultats. A ce titre, la longueur du questionnaire en ligne était susceptible d'en altérer l'acceptabilité, à fortiori auprès d'un public jeune. Le nombre réduit d'échantillons, la taille restreinte des échantillons après appariement et le nombre restreint d'itérations du protocole constituent également des limites méthodologiques importantes, qui enjoignent à considérer avec prudence les résultats de la présente étude et appellent des études confirmatoires à plus grande échelle. De même, on se doit de prendre en compte ici un potentiel effet de Hawthorne, à fortiori dans la mesure où il s'agit d'une recherche-intervention à caractère collaboratif. On peut en outre considérer que l'utilisation de l'hétéro-évaluation des enseignants comme indicateur de performance, relativement aux compétences cibles en maîtrise de la langue et production d'écrit, est susceptible d'occasionner des biais. Enfin, si Moodle Atelier et Moodle en général offrent une grande flexibilité en termes d'usages – et de détournements d'usages –, cette dernière va de pair avec un processus d'appropriation exigeant, susceptible de freiner ou d'empêcher l'adoption de telles innovations pédagonumériques.

CONCLUSION

1 Bilan : réponse à la question de recherche

Pour répondre au défi que représentent les hétérogénéités au collège unique, singulièrement marqué par les inégalités socioscolaires, nous avons envisagé la différenciation pédagogique (DP) comme une approche incontournable mais qui gagnerait, pour être opérationnalisée efficacement, à s'appuyer sur les pédagogies coopératives. Afin de favoriser la réussite de tous les élèves, singulièrement eu égard aux compétences en maîtrise de la langue et production d'écrit en français langue première au cycle 4, nous avons vu que, outre une amélioration des apprentissages coopératifs, la mise en œuvre de la DP implique de favoriser l'autodirection des apprentissages – *i.e.* une autorégulation accrue, une amélioration des compétences métacognitives et une augmentation du sentiment d'efficacité personnelle des élèves. Nous avons alors envisagé une modalité évaluative susceptible de conjuguer ces différents aspects, à savoir la coévaluation formative individualisée, *i.e.* une modalité d'évaluation formative hybride, qui permet de conjuguer hétéro-évaluation par l'enseignant et inter-évaluation entre élèves. Pour instrumenter la coévaluation formative individualisée, le recours à un outil numérique s'est avéré indispensable ; toutefois nous avons vu qu'un tel artefact ne devient véritablement un outil qu'à la faveur d'une genèse instrumentale, susceptible d'occasionner des conflits instrumentaux, eux-mêmes susceptibles d'être régulés par des processus d'ajustement. *In fine*, pour mettre en œuvre la DP, nous nous sommes demandé dans quelle mesure on pourrait favoriser, au moyen d'un dispositif de coévaluation numérique, un apprentissage autodirigé et coopératif du français langue première chez des collégiens de cycle 4.

Pour répondre à cette question de recherche, conformément aux recommandations de la littérature scientifique relativement à l'opérationnalisation de la DP comme objet de recherche, nous avons initié une recherche-intervention de type collaboratif avec trois enseignantes volontaires et avec leurs quatre classes de cycle 4, dans trois collèges différents (péri-urbain, rural, urbain en REP), situés dans les académies de Montpellier et Paris. Nous avons ainsi mis en œuvre un protocole semi-expérimental, co-construit avec les enseignantes partenaires, dans le cadre duquel nous avons pratiqué la coévaluation formative individualisée au moyen de l'artefact numérique Moodle Atelier. Cet artefact a été instrumentalisé à cet effet, au moyen d'une catachrèse. Il a généré pour chaque apprenant, sur la base d'une grille de

coévaluation co-construite en amont par l'enseignante et ses élèves, un score de scripteur ainsi qu'un score de coévaluateur.

Le protocole semi-expérimental impliquait la mise en œuvre d'activités de production écrite ciblant la maîtrise de la langue, en vue d'une coévaluation réalisée en salle informatique de manière hybride, *i.e.* en présentiel et en synchronie, mais par ordinateurs interposés et avec le degré d'asynchronie qu'implique le rythme de travail de chacun. Chacune de ces séances semi-expérimentales était précédée et suivie de la passation d'une échelle psychométrique mesurant le sentiment d'efficacité personnelle des élèves participants, au moyen d'un questionnaire en ligne ; tandis que les scores de scripteurs et de coévaluateurs étaient également recueillis. En outre, des items spécifiques permettaient, dans le questionnaire en ligne, de mesurer le sentiment d'efficacité collective et le degré de conflictualité ressentie relativement au dispositif expérimental de coévaluation individualisée en EIAH. D'autres instruments de recherche ont été mobilisés pour recueillir des données textuelles relatives à la genèse instrumentale du dispositif, *i.e.* des retours écrits sur expérience et des focus groupes avec les élèves participants, ainsi que des entretiens semidirigés et un focus groupe avec les enseignantes partenaires.

Dans le cadre de ce protocole de recherche, les hypothèses opérationnelles que nous nous proposons de tester pour répondre à la question de recherche se rapportaient à l'amélioration des compétences en français (HO 1), à l'amélioration de l'autodirection des apprentissages (HO 2) – et, subséquemment, à l'amélioration de l'autorégulation des apprentissages (HO 2a), à l'amélioration des compétences métacognitives des élèves (HO 2b) et à l'augmentation du sentiment d'efficacité personnelle (HO 2c). Nous avons également testé l'hypothèse HO 3 – *i.e.* « La coévaluation numérique améliore les apprentissages coopératifs » –, ainsi que l'hypothèse HO 4 – *i.e.* « La coévaluation numérique génère des conflits instrumentaux, mais ces derniers peuvent être régulés » –, en examinant successivement cette dialectique conflit/régulation en termes de situations d'apprentissage (HO 4a) et de situations d'enseignement (HO 4b).

Pour tester ces quatre hypothèses, nous avons analysé quantitativement et/ou qualitativement les données numériques et/ou textuelles recueillies dans le cadre du protocole de recherche. Suivant les recommandations de la littérature du champ, eu égard à l'opérationnalisation de la différenciation pédagogique en tant qu'objet de recherche, les résultats de ces analyses, ont été combinés selon le modèle parallèle et convergent de l'approche par méthodes mixtes. Ces résultats nous ont permis de valider les quatre hypothèses opérationnelles. Ainsi, nous avons pu établir que le présent dispositif expérimental de

coévaluation numérique individualisée constitue un moyen de favoriser des apprentissages autodirigés et coopératifs et contribue à l'amélioration des compétences des collégiens de cycle 4 participants en maîtrise de la langue et production d'écrit en français langue première.

En l'occurrence, nous avons pu constater que le dispositif expérimenté ici est propice à la différenciation pédagogique, notamment du fait de son hybridité et des possibilités qu'il offre pour adapter la tâche à la zone proximale de développement de l'apprenant, ainsi que grâce à certaines fonctionnalités spécifiques à l'artéfact numérique. En outre, nous avons vu que, dans l'ensemble, le dispositif expérimental de coévaluation numérique contribue à réduire les inégalités scolaires, tout en favorisant la réussite de tous les élèves.

Nos résultats ont également montré que le dispositif est propice à la mise en œuvre des pédagogies coopératives, en articulant autodirection des apprentissages et coopération entre élèves. La dimension formative du dispositif de coévaluation individualisée a également été mise en évidence, et pourrait faire l'objet d'un approfondissement, notamment grâce à la mise en œuvre d'une boucle formative de rétroaction anonymisée, via une catachrèse ad hoc, pratiquée à même l'artéfact numérique Moodle Atelier.

Surtout, nous avons vu que, si la coévaluation formative individualisée peut être porteuse pour l'amélioration des compétences de cycle 4 en maîtrise de la langue et production d'écrit en français langue première, elle n'est pas envisageable sans le recours à un artéfact numérique tel que Moodle Atelier. Certes, nous avons pu constater que la genèse instrumentale d'un tel artéfact occasionnait des conflits instrumentaux, mais ces derniers peuvent *in fine* être régulés grâce à des processus d'ajustement, selon une « dialectique instrumentale » se rapportant à différentes combinatoire conflictuelles ou synergies régulatrices, et impliquant des relations de causalité réciproque, voire des effets de rétroaction.

Nous avons vu que d'une part, un artéfact numérique comme Moodle Atelier présente l'avantage d'être doté d'une grande plasticité, intrinsèquement liée au fait qu'il s'agit d'un logiciel libre (non-propriétaire) dont les fonctionnalités sont conçues à la faveur des apports constants d'une communauté dynamique de développeurs et d'utilisateurs. Or si cette plasticité se prête bien à des processus d'instrumentalisation virtuellement illimités, et est donc propice au déploiement de la liberté pédagogique des enseignants, elle s'accompagne par ailleurs d'une exigence forte en termes d'appropriation de l'outil, pour apprendre à en maîtriser les rouages et à le paramétrer efficacement. La diffusion d'une pratique comme la coévaluation formative individualisée est donc conditionnée par un développement de l'usage de Moodle qui ne va pas de soi, et ne peut faire l'économie d'un accompagnement institutionnel, notamment au plan de la formation initiale et continue des personnels enseignants. A cet égard, l'importance que

l'Éducation nationale accorde déjà à Moodle constitue une tendance prometteuse, d'une part dans la mesure où, depuis une dizaine d'années déjà elle a fait installer Moodle dans tous les collèges de France – via les environnements numériques de travail (ENT) académiques –, en proposant des formations ad hoc pour les personnels enseignants du secondaire, et d'autre part compte tenu du fait que les Cellules Académiques Recherche, Développement, Innovation et Expérimentation (CARDIE), les délégations académique du numérique pour l'éducation (DANE et DRANE) et les corps d'inspection de Lettres (*e.g.* GIPTIC Lettres) se sont déjà largement emparés de cet enjeu. A ce titre, l'accueil très favorable qui a été réservé à la présente recherche par les différents partenaires précités, ainsi que la qualité de l'accompagnement qu'ils ont proposé au chercheur-intervenant pour lui faciliter l'accès et l'usage de Moodle Atelier, en témoignent largement.

Pour finir, eu égard à la question très concrète qui traverse l'ensemble de cette recherche, à savoir comment réduire les inégalités socioscolaires au collège, et y favoriser au contraire la réussite de tous les élèves au moyen de la différenciation pédagogique, ce dispositif expérimental de coévaluation numérique représente un succès en demi-teinte. Il semble néanmoins que nos résultats puissent être considérés comme encourageants, compte tenu de l'effet de levier que la métacognition joue sur les apprentissages en général (Efklides, 2006; Pintrich *et al.*, 2000), à fortiori en lien avec le concept d'autodirection (Carré, 2010; Efklides *et al.*, 2017), et en prenant en compte les dimensions socio-affectives – voire coopératives – que toute situation d'enseignement/apprentissage met en jeu (Efklides, 2011; Zimmerman, 1995). A l'égard de ces concepts à la fois connexes et décisifs que nous avons cherché à opérationnaliser ici – et dont le sentiment d'efficacité personnelle (Bandura, 2007; Berger & Büchel, 2012; Carré, 2004) constitue une des clés de voûte –, cette recherche contribuera peut-être à montrer que la coévaluation numérique est une voie qui gagne à être explorée plus avant.

A ce titre, puisqu'il s'agit selon nous de *mettre le numérique au service de la différenciation pédagogique* – cet objet complexe qui, après presque quarante ans de recherches, « résiste encore à sa propre opérationnalisation » (Forget, 2017b, p. 49) –, l'approche instrumentale, développée dans le prolongement de (Rabardel, 1995) par des auteurs comme (Marquet, 2005b, 2005a, 2011; Zeller & Marquet, 2021b), semble un cadre théorique à même d'accompagner efficacement à la fois la recherche *et* les usages des praticiens – notamment dans le cadre d'études collaboratives et/ou en employant des méthodes mixtes (Forget, 2017b, p. 50).

2 Perspectives pour la recherche et l'enseignement

Pour pratiquer la coévaluation numérique, plutôt que de recourir à des solutions propriétaires, certes souvent bien plus ergonomiques, avec des interfaces très intuitives – mais qui tendent à contraindre les schèmes d'utilisation (*i.e.* instrumentation) –, il nous semble que des logiciels libres comme Moodle sont davantage propices à l'instrumentalisation que requiert la liberté pédagogique des enseignants, à fortiori compte tenu des enjeux de *souveraineté* que le développement et l'utilisation de logiciels libres constituent pour l'Ecole de la République. Il s'agit cependant d'une voie exigeante : le fait que Moodle soit installé dans tous les collèges de France représente déjà, en soi, un atout majeur ; toutefois un accompagnement accru des enseignants souhaitant s'approprier cet outil semble nécessaire. Dans cette perspective, ce projet a montré que les enseignants de collège peuvent s'approprier Moodle en général, et Moodle Atelier en particulier, de manière très féconde, voire contribuer à faire évoluer Moodle en s'associant à la communauté dynamique qui gravite autour de cette plateforme logicielle. A ce titre, l'enseignante C a par exemple suggéré que, dans Moodle Atelier, au moment de la coévaluation proprement dite, il faudrait pouvoir annoter le texte évalué au fur et à mesure ; car ensuite pour remplir le formulaire d'évaluation ou pour rédiger l'écrit commentatif (appréciation), il n'est pas toujours évident de se souvenir des faits de langue concernés – et dont l'évaluateur a besoin pour fonder ou argumenter son évaluation. Actuellement on est en effet obligé, pour ce faire, de faire défiler l'écran pour effectuer des allers-retours entre le texte source et la grille de coévaluation, ou entre le texte source et l'encart destiné à accueillir l'écrit commentatif. D'un point de vue cognitif, cet aspect ergonomique pose problème. Cette préconisation de l'enseignante C est un exemple illustrant le type de contributions concrètes, fondées sur l'usage – *i.e.* expérience d'utilisateur ou UX (*user experience*) –, que des enseignants de collège pourraient apporter à la communauté Moodle, pour participer à l'évolution de l'outil, en faveur d'une ergonomie sans cesse améliorée.

Récemment, la participation de nos partenaires de la CARDIE, de la DANE et du GIPTIC Lettres du rectorat d'académie de Paris au Moodle Moot qui s'est tenu du 5 au 7 juillet à l'université de technologie de Troyes (UTT) montre l'importance que l'Education nationale accorde à un outil comme Moodle. De telles initiatives constituent un autre signe encourageant, en faveur du développement des usages de Moodle parmi les enseignants du secondaire en général, et potentiellement parmi les enseignants de collège en particulier. De même, alors que

la nouvelle plateforme Moodle ELEA nationale est en train d'être déployée, le GIPTIC Lettres de l'académie de Paris mène des actions ad hoc de formation continue, et promeut ainsi le développement d'une communauté active de praticiens experts qui, discipline par discipline, œuvrent auprès de leurs collègues à la diffusion des innovations pédagogiques et des bonnes pratiques liées aux usages de Moodle.

A cet égard, des projets comme la présente recherche collaborative paraissent à même de faire évoluer Moodle en général et Moodle Atelier en particulier, en offrant à des chercheurs-intervenants et à des enseignants partenaires l'opportunité de co-construire des connaissances d'autant plus à même de contribuer au développement – notamment ergonomique – de l'artéfact numérique, qu'elles sont issues, épistémologiquement, de ces « échanges, migrations et métissages de savoirs, de natures et de statuts différents » (Marcel, 2016b, p. 31). Singulièrement, de tels projets de recherche gagneraient à s'inscrire dans la dynamique communautaire dédiée au développement de Moodle, et en partenariat avec une variété d'acteurs : chercheurs, enseignants, élèves et partenaires institutionnels tels que ceux précités – selon l'approche de la *community-based research* (Jason & Glenwick, 2016), voire de la *community-based participatory research* (Campano *et al.*, 2020; Zuber-Skerritt *et al.*, 2020).

Par ailleurs, une fonctionnalité déjà instrumentée dans l'artéfact numérique, et destinée initialement à l'inter-évaluation, pourrait tout à fait être mobilisée pour la coévaluation. En l'occurrence, dans une logique de retravail des productions écrites à la lumière des écrits commentatifs reçus, les élèves peuvent se voir proposer des « Travaux publiés », anonymisés ou non : autant de points d'appui, issus du travail de « pairs experts », auxquels ils pourraient néanmoins s'identifier pour y puiser les ressources nécessaires à une montée en compétences pour leurs propres écrits futurs. Les élèves qui le souhaitent pourraient en outre retravailler leur texte avant de le soumettre à une autre évaluation – formative ou sommative – et/ou donner leur autorisation pour que leurs travaux soient publiés, non pas à titre d'exemples émanant de pairs experts, mais en tant que travaux dans lesquels telle ou telle compétence pourrait être retravaillée. Là encore, une telle publication pourrait se faire de manière anonymisée ou non : au choix de l'élève. De tels scénarios, fondés sur le volontariat de l'élève et adaptables à une variété de degrés de maîtrise des compétences cibles, ont spontanément été mis en place par l'enseignante C avec sa « classe coop' », dans la mesure où il s'agissait pour ses élèves et elles de pratiques déjà bien établies. Au moyen de la fonctionnalité « Travaux publiés » – déjà instrumentée dans l'artéfact numérique Moodle Atelier, donc facilement mobilisable –, on pourrait ainsi approfondir la différenciation pédagogique au sein du dispositif de coévaluation numérique.

A ce titre, les perspectives esquissées pour approfondir la catachrèse mise en œuvre dans la présente recherche, en vue d'une boucle formative de rétroaction anonymisée, mériteraient de faire l'objet d'une recherche future. A fortiori, une telle recherche pourrait être entreprise selon le paradigme de la *community-based participatory research*. En effet, compte tenu des difficultés techniques afférentes, on gagnerait à y associer étroitement la communauté des développeurs et usagers de Moodle, ainsi que des partenaires institutionnels dont on sait qu'ils sont déjà très mobilisés sur ces questions, *e.g.* Ministère de l'Education nationale, rectorats d'académie, corps d'Inspection, CARDIE, DANE/DRANE, GIPTIC Lettres. Pour rappel, la catachrèse en question, suggérée par les élèves eux-mêmes, constitue une piste de régulation instrumentale en réponse à plusieurs conflits instrumentaux didactiques et pédagogiques rencontrés – *e.g.* le caractère insuffisamment explicite de certains écrits commentatifs de pairs ou le sentiment d'injustice ressenti à l'égard de certaines productions écrites jugées méritoires, mais difficilement valorisables car elles ne correspondent pas suffisamment aux critères de coévaluation. Concrètement, cette boucle formative de rétroaction anonymisée permettrait aux élèves coévalués, aux pairs coévaluateurs et à l'enseignant d'interagir au sein de l'artéfact numérique, de manière synchrone ou asynchrone, et tout en conservant – pour les élèves – l'usage de la fonctionnalité « anonymisation en double aveugle ». Signalons que cette proposition de catachrèse directement issue des utilisateurs fait écho à un scénario pédagonumérique alternatif que nous avons envisagé lors de la phase exploratoire de la présente recherche, en nous référant aux travaux de (Connac & Cieutat, 2021 ; Reynaud, 2022). Une telle catachrèse pourrait s'appuyer en partie sur la fonctionnalité « Travaux publiés », et serait propice à une mise en œuvre accrue de la différenciation pédagogique dans le cadre du dispositif expérimental de coévaluation numérique, dans la mesure où elle permettrait à chaque élève d'effectuer, en coopérant avec ses pairs et avec l'enseignant, un parcours d'apprentissage personnalisé, adapté à ses besoins réels, pour parvenir *in fine* à une maîtrise satisfaisante des attendus didactiques (Forget, 2017b). La Figure 91 propose une représentation schématique de ce à quoi cette catachrèse « Boucle de rétroaction formative anonymisée » pourrait s'apparenter.

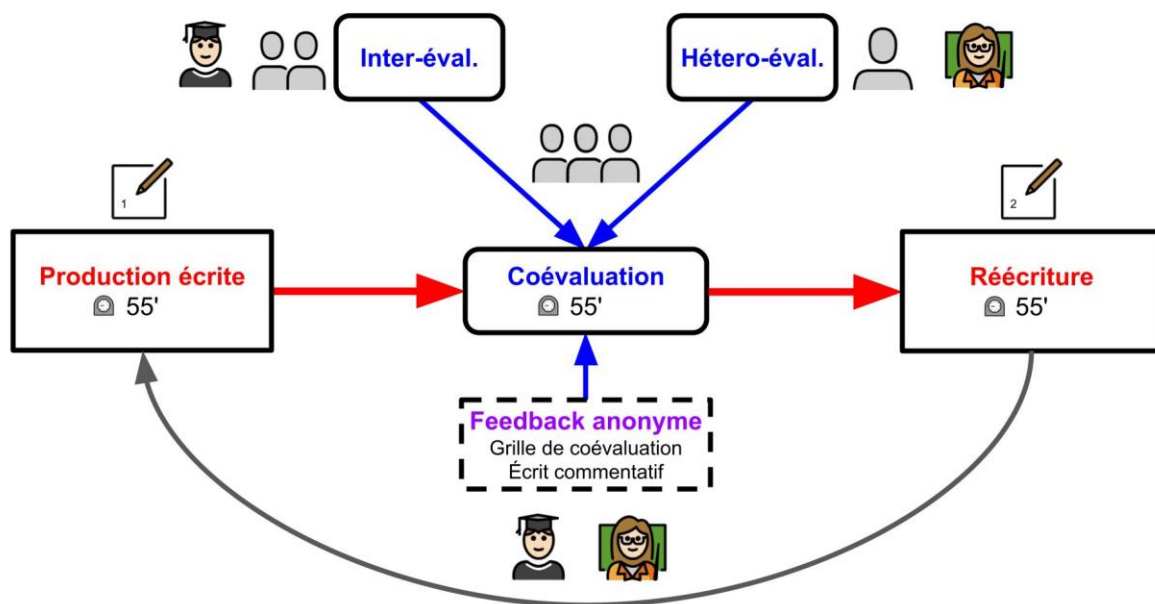


Figure 91 – Proposition de représentation schématique pour la catachrèse « Boucle de rétroaction formative anonymisée »

La présente recherche peut par ailleurs nourrir d'autres études ciblant plus spécifiquement les écrits commentatifs, à l'instar des travaux de (Brunel, 2018; Brunel & Guérin-Callebout, 2018; Taous, 2020), ou approfondir le travail sur les productions écrites initiales, en prenant en compte la spécificité de leur production sur traitement de texte (Doquet-Lacoste, 2006). Un certain nombre d'aspects en lien avec la didactique du français pourraient effectivement être approfondis, notamment pour une observation plus fine de l'évolution de la maîtrise des compétences cibles, en examinant de manière systématique les productions d'élèves (Jacques & Poibeau, 2010; Jacques & Rinck, 2017), via une approche textométrique avec un logiciel comme TXM par exemple, pour y analyser des faits de langue précis, en prenant en compte la spécificité de ces écrits produits sur traitement de texte (Doquet-Lacoste, 2006). Une telle approche jouerait un rôle complémentaire vis-à-vis des autres approches analytiques développées dans cette recherche, notamment dans la perspective d'une approche par méthodes mixtes – *i.e.* en combinant les résultats liés à d'autres indicateurs avec des résultats fondés sur la mobilisation de ressources textométriques pour l'analyse linguistique des écrits d'élèves. De surcroit, relativement aux métacompétences afférentes, les écrits commentatifs des élèves coévaluateurs pourraient faire l'objet d'une analyse systématique, en lien avec les écrits de scripteurs auxquels ils se rapportent, afin de mieux comprendre la manière dont les élèves parviennent à mobiliser ou non les compétences et métacompétences nécessaires à la formulation d'un écrit commentatif argumenté à destination d'un pair, à l'instar des travaux de (Brunel, 2018; Brunel & Guérin-Callebout, 2018; Taous, 2020).

En termes psychopédagogiques et ergonomiques, on pourrait investiguer plus avant la manière dont le dispositif expérimental de coévaluation numérique constitue une aide à l'autodirection des apprentissages (Mailles Viard Metz, 2015b; Mailles Viard Metz *et al.*, 2014), en s'intéressant tout particulièrement à la dialectique féconde entre outils, étayages et autonomie (Duthoit *et al.*, 2012; Mailles Viard Metz, 2015a) et en interrogeant plus avant les effets du dispositif expérimental sur les sentiments d'efficacité personnelle et collective des participants (Bellet *et al.*, 2019).

Sur le plan épistémologique, il peut être utile d'approfondir la démarche collaborative qui est à l'œuvre ici, et ses effets rétroactifs sur les pratiques (Torterat *et al.*, 2019, 2020), en analysant les échanges entre chercheur-intervenant et enseignants partenaires et en s'appuyant notamment, pour ce faire, sur les documents échangés (progressions annuelles, tableaux synoptiques de séquences de cours...), ainsi que sur les différentes traces de ces échanges, notamment les courriels, considérés comme des écrits de travail (Torterat, 2015).

Surtout, sur la base de la « dialectique instrumentale » mise en évidence entre artéfacts conflictuels d'une part, et artéfacts régulateurs d'autre part (Zeller & Marquet, 2021b), et en lien avec la typologie relative au degré de conflictualité des phénomènes observés (Marquet, 2005b, 2011), on pourrait également entreprendre un travail de modélisation plus systématique (Marquet, 2003), pour tenter de faire émerger un modèle intégrateur de ces différentes combinatoires artéfactuelles et synergies régulatrices. Un tel modèle pourrait notamment permettre d'envisager la dynamique conflit/régulation comme un continuum, opérationnalisable en termes de prévention des conflits instrumentaux et de mise en œuvre de stratégies régulatrices.

En somme, eu égard à la différenciation pédagogique, ces différentes perspectives pourraient s'avérer porteuses à la fois pour la recherche et pour les usages du numérique en classe, à plus forte raison s'agissant d'un artéfact à la fois riche de potentialités et exigeant comme Moodle Atelier.

REFERENCES BIBLIOGRAPHIQUES

- Ahr, S., Czarny, N., Féki, H., & Moinard, P. (2011). L'espace numérique collaboratif : un lieu pour former et se former, pour enseigner et apprendre. *Éla. Études de linguistique appliquée*, n°160(4), 465-477. <https://doi.org/10.3917/ela.160.0465>
- Ajaja, O. P., & Eravwoke, O. U. (2010). Effects of cooperative learning strategy on junior secondary school students achievement in integrated science. *The Electronic Journal for Research in Science & Mathematics Education*, 14(1).
- Algan, Y., Huillery, É., & Prost, C. (2018). Confiance, coopération et autonomie : pour une école du XXIe siècle. *Notes du conseil d'analyse économique*, 48(3), 1. <https://doi.org/10.3917/NCAE.048.0001>
- Ali, Z., & Bhaskar, Sb. (2016). Basic statistical tools in research and data analysis. *Indian Journal of Anaesthesia*, 60(9), 662-669. <https://doi.org/10.4103/0019-5049.190623>
- Allal, L. (2000). Metacognitive regulation of writing in the classroom. *Metalinguistic activity in learning to write*, 6, 145-166.
- Allal, L. (2001). Situated cognition and learning: From conceptual frameworks to classroom investigations. *Schweizerische Zeitschrift für Bildungswissenschaften*, 23(3), 407-422.
- Allal, L. (2017). Comment former les enseignants à l'évaluation au service de la différenciation pédagogique ? In Cnesco (Éd.), *Notes remises dans le cadre de la conférence de consensus du Cnesco et de l'Ifé/Ens de Lyon « Différenciation pédagogique : comment adapter l'enseignement pour la réussite de tous les élèves ? »* (p. 167-176).
- Allal, L. (2018). Peut-on intégrer l'enseignement de l'orthographe dans la production de textes? *Ecrire et rédiger: Comment accompagner les élèves dans leurs apprentissages?*
- Allal, L. (2020). Assessment and the co-regulation of learning in the classroom. *Assessment in Education: Principles, Policy & Practice*, 27(4), 332-349. <https://doi.org/10.1080/0969594x.2019.1609411>
- Allal, L. (2021). Involving primary school students in the co-construction of formative assessment in support of writing. *Assessment in Education: Principles, Policy and Practice*, 28(5-6), 584-601. <https://doi.org/10.1080/0969594X.2021.1951164>
- Anadón, M. (2019a). Les méthodes mixtes : implications pour la recherche « dite » qualitative. *Recherches qualitatives*, 38(1), 105-123. <https://doi.org/10.7202/1059650ar>
- Anadón, M. (2019b). Les méthodes mixtes : implications pour la recherche « dite » qualitative. *Recherches qualitatives*, 38(1), 105. <https://doi.org/10.7202/1059650ar>

- Arroyave Buriticá, M. C., Giraldo Prieto, M. E., & González Ramírez, M. (2007). Model for the identification of work competences and definition of performance criteria in public health of the nursing and medicine professionals in the department of Caldas. *Hacia la Promoción de la Salud*, 12(1), 179-192.
- Ary, D., Jacobs, L. C., Sorensen, C., & Razavieh, A. (2010). Introduction to research in education eight edition. *Wadsworth: Cengage Learning*.
- Astolfi, J.-P. (1995). Essor des didactiques et des apprentissages scolaires. *Éducatives*, 6, 47.
- Attali, J. (2008). Rapport de la Commission pour la libération de la croissance française. *XO Editions, La documentation française*, 16, 47-52.
- Azorín, C. (2020). Beyond COVID-19 supernova. Is another education coming? *Journal of Professional Capital and Community*, 5(3/4), 381-390. <https://doi.org/10.1108/JPCC-05-2020-0019>
- Baker, M. (2008). Formes et processus de la résolution coopérative de problèmes: des savoirs aux pratiques éducatives. *Vers des apprentissages en coopération: rencontres et perspectives*, 107-130.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Bandura, A. (2000). Exercise of human agency through collective efficacy. *Current Directions in Psychological Science*, 9(3), 75-78. <https://doi.org/10.1111/1467-8721.00064>
- Bandura, A. (2002). Growing primacy of human agency in adaptation and change in the electronic era. *European psychologist*, 7(1), 2.
- Bandura, A. (2006). Guide For Constructing Self-Efficacy Scales. In T. Urdan & F. Pajares (Éds.), *Self-Efficacy Beliefs of Adolescents* (p. 307-337).
- Bandura, A. (2007). Auto-efficacité: le sentiment d'efficacité personnelle (2e éd.). In *Bruxelles: De Boeck*. De Boeck.
- Bandura, A. (2009). La théorie sociale cognitive : une perspective agentique. In *Traité de psychologie de la motivation* (p. 15). Dunod. <https://doi.org/10.3917/dunod.carre.2009.01.0015>
- Bandura, A. (2011). Social cognitive theory. In *Handbook of Theories of Social Psychology* (Vol. 1, p. 349-373).
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Multifaceted Impact of Self-Efficacy Beliefs on Academic Functioning. In *Multifaceted Impact of Self-Efficacy Beliefs on Academic Functioning*. *CHILD DEVELOPMENT* (Vol. 67).

- Barth, B.-M. (2013). *Élève chercheur, enseignant médiateur: donner du sens aux savoirs*. Retz.
<https://www-numeriquepremium-com.ezpupv.scdi-montpellier.fr/content/books/9782725631585>
- Barth, B.-M. (2015). *Le savoir en construction*. Retz.
- Baudrit, A. (2007). *Relations d'aide entre élèves à l'école*. De Boeck Supérieur.
<https://doi.org/10.3917/dbu.baudr.2007.02>
- Beer-Toker, M., Huel, C., & Richer, R. (2009). La chaise de l'auteur et le traitement de texte : leurs effets sur le processus d'écriture d'élèves en difficulté. *Revue des sciences de l'éducation*, 17(3), 465-484. <https://doi.org/10.7202/900710ar>
- Bellet, P. (2022). *De l'efficacité collaborative à l'intelligence collective : mise en œuvre instrumentale en situation d'apprentissage*. Université Paul-Valéry Montpellier (UPVM).
- Bellet, P., Vendeville, N., & Mailles Viard Metz, S. (2019). Les effets du travail de groupe sur les sentiments d'efficacité personnelle et collective en contexte d'apprentissage. *Canadian Journal of Education/Revue canadienne de l'éducation*, 42(4), 992-1021.
<https://www.jstor.org/stable/26954665>
- Berger, J.-L., & Büchel, F. (2012). Métacognition et croyances motivationnelles : un mariage de raison. *Revue française de pédagogie*, 179, 95-128. <https://doi.org/10.4000/rfp.3705>
- Bertaux, D. (2003). Les récits de vie. In *Paris: Nathan université*. Nathan Université.
- Bieber, T., & Martens, K. (2011). The OECD PISA Study as a Soft Power in Education? Lessons from Switzerland and the US. *European Journal of Education*, 46(1), 101-116.
<https://doi.org/10.1111/j.1465-3435.2010.01462.x>
- Black, P., Harrison, C., Lee, C., & Wiliam, D. (2005). Changer l'enseignement grâce à l'évaluation formative : recherche et pratique. The King's Medway Oxfordshire formative assessment project. In CERI: Centre pour la recherche et l'innovation dans l'enseignement (Éd.), *L'évaluation formative : pour un meilleur apprentissage dans les classes secondaires* (p. 245-264). OCDE.
<https://books.google.fr/books?id=hOiaAwAAQBAJ&lpg=PA265&ots=GGsBr-9Bff&lr&pg=PA245#v=onepage&q&f=false>
- Blaikie, N. (2003). *Analyzing quantitative data: From description to explanation*. Sage.
- Blau, I., & Shamir-Inbal, T. (2018). Digital technologies for promoting “student voice” and co-creating learning experience in an academic course. *Instructional Science*, 46(2), 315-336.
<https://doi.org/10.1007/s11251-017-9436-y>

- Blettery, B., Ebrahim, L., Honnart, D., & Aube, H. (1996). Les échelles de mesure de la douleur dans un service d'accueil des urgences. *Réanimation Urgences*, 5(6), 691-697. [https://doi.org/10.1016/S1164-6756\(05\)80594-8](https://doi.org/10.1016/S1164-6756(05)80594-8)
- Bogo, M., Katz, E., Regehr, C., Logie, C., Mylopoulos, M., & Tufford, L. (2013). Toward Understanding Meta-Competence: An Analysis of Students' Reflection on their Simulated Interviews. *Social Work Education*, 32(2), 259-273. <https://doi.org/10.1080/02615479.2012.738662>
- Bolzan, L. M., Otero, W. R. I., & Botelho, R. M. (2022). Self-assessment and hetero-assessment: practical approach in higher education. *Estudos em Avaliação Educacional*, 33.
- Boone, H. N., & Boone, D. A. (2012). Analyzing likert data. *Journal of extension*, 50(2), 1-5.
- Boris, H., Campbell, C., Cavanagh, M., Petocz, P., & Kelly, N. (2013). Technological pedagogical content knowledge of secondary mathematics teachers. *Contemporary Issues in Technology and Teacher Education*, 13(1), 22-40.
- Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, 31(4), 399-413. <https://doi.org/10.1080/02602930600679050>
- Bourdieu, P., & Passeron, J.-C. (1970). *La reproduction: éléments pour une théorie du système d'enseignement*. Les éditions de Minuit.
- Bouziane, A., & Zyad, H. (2018a). The Impact of Self and Peer Assessment on L2 Writing: The Case of Moodle Workshops. In *Assessing EFL Writing in the 21st Century Arab World* (p. 111-135). Springer International Publishing. https://doi.org/10.1007/978-3-319-64104-1_5
- Bouziane, A., & Zyad, H. (2018b). The impact of self and peer assessment on L2 writing: The case of moodle workshops. In *Assessing EFL Writing in the 21st Century Arab World: Revealing the Unknown* (p. 111-135). Springer International Publishing. https://doi.org/10.1007/978-3-319-64104-1_5
- Bouzidi, L., & Jaillet, A. (2009). Can online peer assessment be trusted? *Journal of Educational Technology & Society*, 12(4), 257-268.
- Braun, H. (2003). Assessment and Technology. In *Optimising New Modes of Assessment: In Search of Qualities and Standards* (p. 267-288). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-48125-1_11

- Brookhart, S. M. (2001). Successful Students' Formative and Summative Uses of Assessment Information. *Assessment in Education: Principles, Policy & Practice*, 8(2), 153-169. <https://doi.org/10.1080/09695940123775>
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.
- Brunin, P., Morois, G., Lavit-Bérard, C., & de Khovrine, N. (2023, mars). Une recherche collaborative sur l'évaluation coopérative. *Les Cahiers pédagogiques*, 32-33.
- Brunel, M. (2018). Les écrits de fanfiction dans la classe. *Le français aujourd'hui*, N° 200(1), 31-42. <https://doi.org/10.3917/lfa.200.0031>
- Brunel, M., & Guérin-Callebout, C. (2018). « Écrire dans » : Écriture littéraire sur écran. Présentation d'une expérimentation en classe de 3^e année du secondaire. *Revue de recherches en littératie médiatique multimodale*, 3. <https://doi.org/10.7202/1047128ar>
- Bruner, J. S. (1983). *Le développement de l'enfant: savoir faire, savoir dire*. Presses universitaires de France.
- Büchel, F. P. (1991). How are learning activities initiated? An epistemological view of mediated learning experience. In R. Feuerstein, P. S. Klein, & A. J. Tannenbaum (Éds.), *Mediated learning experience (MLE): Theoretical, psychological and learning implications* (p. 53-70). Freund Publishing House.
- Bui, T. T. H. (2007). *Le sentiment d'autonomie de l'enfant par rapport à l'école: analyse comparée en France et au Vietnam*. Paris 10 Nanterre.
- Buisson-Fenet, H., & Marx, L. (2022). Enseigner et apprendre à l'épreuve de la « continuité pédagogique ». *Recherches en éducation*, 48. <https://doi.org/10.4000/ree.11125>
- Burns, R. (1971). Methods for individualizing instruction. *Educational technology*, 11(6), 55-56.
- Burrick, D. (2010). Une épistémologie du récit de vie. *Recherche qualitative*, 8, 7-36.
- Campano, G., Ghiso, M. P., Badaki, O., & Kannan, C. (2020). Agency as collectivity: Community-based research for educational equity. *Theory Into Practice*, 59(2), 223-233. <https://doi.org/10.1080/00405841.2019.1705107>
- Cardinet, J. (1988). Évaluation scolaire et mesure. *Pédagogies en développement. Recueils*.
- Caron, P.-A. (2020). Ingénierie dispositif et enseignement à distance au temps de la COVID 19. *Distances et médiations des savoirs*, 30. <https://doi.org/10.4000/dms.5211>
- Carré, P. (2003). Aux sources de l'autoformation, l'éducation nouvelle ? In H. Lenoir & E. M. Lipiansky (Éds.), *Recherches et innovations en formation* (L'Harmattan, p. 99-112).

- Carré, P. (2004). Bandura : une psychologie pour le XXIe siècle ? *Savoirs, Hors série*(5), 9. <https://doi.org/10.3917/savo.hs01.0009>
- Carré, P. (2010). Chapitre 3. L'autodirection des apprentissages. In *L'autoformation* (p. 117-169). Presses Universitaires de France. <https://doi.org/10.3917/puf.carre.2010.02.0117>
- Carré, P., Moisan, A., Poisson, D., & Moisan, A. (1997). *L'autoformation: psychopédagogie, ingénierie, sociologie*. Presses universitaires de France.
- Cavalier, J. C., Klein, J. D., & Cavalier, F. J. (1995). Effects of cooperative learning on performance, attitude, and group behaviors in a technical team environment. *Educational Technology Research and Development*, 43(3), 61-72. <https://doi.org/10.1007/BF02300456>
- Cayouette-Rembrière, J., & Moulin, L. (2019). Comment évoluent les inégalités de performances scolaires au collège ? Un suivi longitudinal des élèves entre la 6e et la 3e. *Population, Vol. 74*(4), 551-586. <https://doi.org/10.3917/popu.1904.0551>
- Chapelain, B. (2017). La participation dans les écritures créatives en réseaux : de la réception à la production. *Le français aujourd'hui, N° 196*(1), 45-56. <https://doi.org/10.3917/lfa.196.0045>
- Chartier, P. (2013). *La fiabilité des évaluations scolaires* (p. 119-135). <https://doi.org/10.3917/edagri.marce.2013.01.0119>
- Chen, C. hsiu. (2010). The implementation and evaluation of a mobile self- and peer-assessment system. *Computers and Education*, 55(1), 229-236. <https://doi.org/10.1016/j.compedu.2010.01.008>
- Chilotti, P. (2013). L'espace tutorial: de l'hétéro-évaluation à l'autoévaluation. *Actes du 25ème colloque de l'ADMEE-Europe*.
- Cnesco. (2016). *Comment l'école amplifie-t-elle les inégalités sociales et migratoires ? Dossier de synthèse*.
- Cnesco. (2017). *Différenciation pédagogique : comment adapter l'enseignement pour la réussite de tous les élèves ? Dossier de synthèse*.
- Cohen, L., Manion, L., & Morrison, K. (2002). *Research Methods in Education*. Routledge. <https://doi.org/10.4324/9780203224342>
- Colognesi, S., & Van Nieuwenhoven, C. (2017). Le processus de coévaluation entre superviseurs et étudiants en formation initiale des enseignants du primaire. *Canadian Journal of Education/Revue canadienne de l'éducation*, 40(2), 1-27.

- Connac, S. (2016). Autonomie, responsabilité et coopération : ce qu'en disent les élèves utilisant un plan de travail. *Éducation et socialisation*, 41. <https://doi.org/10.4000/edso.1725>
- Connac, S. (2017a). *Enseigner sans exclure: la pédagogie du colibri*. ESF.
- Connac, S. (2017b). *La coopération entre élèves*. Canopé éditions.
- Connac, S. (2018). Ce que disent des élèves sur les classes coopératives en collège et lycée. *Tréma*, 50. <https://doi.org/10.4000/trema.4265>
- Connac, S. (2022). *La personnalisation des apprentissages: agir face à l'hétérogénéité, à l'école et au collège*. ESF sciences humaines.
- Connac, S., & Cieutat, P. (2021). *Coopération et évaluation. Pour ne décourager aucun élève*. Chronique sociale.
- Conseil d'évaluation de l'École (CEE). (2022). *Rapport annuel du conseil d'évaluation de l'École*.
- Cosnefroy, L. (2010). Se mettre au travail et y rester : les tourments de l'autorégulation. *Revue française de pédagogie*, 170, 5-15. <https://doi.org/10.4000/rfp.1388>
- Cosnefroy, L. (2011). L'apprentissage autorégulé : perspectives en formation d'adultes. *Savoirs*, n° 23(2), 9-50. <https://doi.org/10.3917/savo.023.0009>
- Cosnefroy, L., & Fenouillet, F. (2009). Motivation et apprentissages scolaires. P. Carré & F. Fenouillet (Éd.), *Traité de psychologie de la motivation*, 127-146.
- Cosnefroy, L., & Jézégou, A. (2013). Les processus d'autorégulation collective et individuelle au cours d'un apprentissage par projet. *Revue internationale de pédagogie de l'enseignement supérieur*, 29(2). <https://doi.org/10.4000/ripes.744>
- Crahay, M. (2013). *L'école peut-elle être juste et efficace? De l'égalité des chances à l'égalité des acquis*. De Boeck Supérieur.
- Crahay, M., & Delhaxhe, A. (2003). L'enseignement secondaire inférieur: entre culture de l'intégration et culture de la différenciation. *Rapport PIREF*.
- Crahay, M., Hindryckx, G., & Lebe, M. (2001). Analyse des interactions entre enfants en situation de tutorat portant sur des problèmes mathématiques de type multiplicatif. *Revue française de pédagogie*, 136(1), 133-145. <https://doi.org/10.3406/rfp.2001.2833>
- Creswell, J. W., & Creswell, J. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications Thousand Oaks, CA.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. <https://doi.org/10.1007/BF02310555>

- Croze, E. (2021). Covid-19 et passage éclair au distanciel pour les enseignants de langues vivantes du secondaire : une expérience renvoyant à l'irréductible présence en classe de langue. *Distances et médiations des savoirs*, 33. <https://doi.org/10.4000/dms.6134>
- Daniel, S. J. (2020). Education and the COVID-19 pandemic. *Prospects*, 49(1-2), 91-96. <https://doi.org/10.1007/s11125-020-09464-3>
- Darwanto, B. A., Surjono, H. D., & Ciptaningrum, D. S. (2020). The Impact of Digitally Regulated Peer Corrective Feedback (CF) on Students' English Writing. *Universal Journal of Educational Research*, 8(11), 5388-5394. <https://doi.org/10.13189/ujer.2020.081141>
- De Ketele, J.-M. (1993). L'évaluation conjuguée en paradigmes. *Revue française de pédagogie*, 103(1), 59-80. <https://doi.org/10.3406/rfp.1993.1298>
- De Landsheere, G. (1971). Evaluation continue et examens: précis de docimologie. *Education 2000*.
- Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology*, 49(3), 182-185. <https://doi.org/10.1037/a0012801>
- Deeley, S. J. (2014). Summative co-assessment: A deep learning approach to enhancing employability skills and attributes. *Active Learning in Higher Education*, 15(1), 39-51. <https://doi.org/10.1177/1469787413514649>
- Degenne, A., Vergès, P., & Verges, P. (1973). Introduction à l'analyse de similitude. *Revue Française de Sociologie*, 14(4), 471. <https://doi.org/10.2307/3320247>
- Dekhinet, R., Topping, K., Duran, D., & Blanch, S. (2008). Let Me Learn with My Peers Online!: Foreign Language Learning Through Reciprocal Peer Tutoring. *Innovate*, 4(3). <http://www.innovateonline.info/index.php?view=article&id=479>
- Delahaye, J.-P. (2015). Grande pauvreté et réussite scolaire. In *Le choix de la solidarité pour la réussite de tous*.
- Delbrassine, D. (2018). L'usage du traitement de texte pour les tâches de production écrite au lycée. *Le français aujourd'hui*, N° 200(1), 97-116. <https://doi.org/10.3917/lfa.200.0097>
- Demirel, M. (2009). Lifelong learning and schools in the twenty-first century. *Procedia - Social and Behavioral Sciences*, 1(1), 1709-1716. <https://doi.org/10.1016/j.sbspro.2009.01.303>
- Devenney, R. (1989). How ESL Teachers and Peers Evaluate and Respond to Student Writing. *RELC Journal*, 20(1), 77-90. <https://doi.org/10.1177/003368828902000106>

- Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively? *Educational Research Review*, 3(2), 101-129. <https://doi.org/10.1016/j.edurev.2008.02.003>
- Direction de l'évaluation de la prospective et de la performance (DEPP). (2021). *L'état de l'École*.
- Direction générale de l'enseignement scolaire. (2020). Programme d'enseignement du cycle des approfondissements (cycle 4). In *D'après le BOEN n° 31 du 30 juillet 2020*. Ministère de l'Éducation nationale.
- Divjak, B., & Maretić, M. M. (2017). Learning Analytics for Peer-assessment: (Dis)advantages, Reliability and Implementation. In *JIOS* (Vol. 41, Numéro 1).
- Dochy, F. (2001). A new assessment era: different needs, new challenges. *Learning and Instruction*, 10, 11-20. [https://doi.org/10.1016/S0959-4752\(00\)00022-0](https://doi.org/10.1016/S0959-4752(00)00022-0)
- Dochy, F., Segers, M., & Sluijsmans, D. (1999). The Use of Self-, Peer and Co-assessment in Higher Education: A review. *Studies in Higher Education*, 24(3), 331-350. <https://doi.org/10.1080/03075079912331379935>
- Doquet-Lacoste, C. (2006). L'objet insaisissable : l'écriture sur traitement de texte. *Genesis*, 27(1), 35-44. <https://doi.org/10.3406/item.2006.1385>
- Dortier, J.-F. (2014). Peut-on apprendre à apprendre ? Métacognition : un tour d'horizon. In *Apprendre* (p. 25-32). Éditions Sciences Humaines. <https://doi.org/10.3917/sh.bedin.2014.01.0025>
- Dunn, T. J., Baguley, T., & Brunsten, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399-412. <https://doi.org/10.1111/bjop.12046>
- Duru-Bellat, M. (2019). Evaluations, mesures ou classements ? A propos des enquêtes PISA. *Revue Française de Linguistique Appliquée*, 24(1), 7-19. <https://doi.org/10.3917/rfla.241.0007>
- Duru-Bellat, M., & Kieffer, A. (1999). Evaluer la démocratisation de l'enseignement : la situation française à l'épreuve des comparaisons internationales. *Revue française de pédagogie*, 127(1), 49-61. <https://doi.org/10.3406/rfp.1999.1084>
- Duru-Bellat, M., & Kieffer, A. (2000). La démocratisation de l'enseignement en France: Polemiques autour d'une question d'actualité. *Population (French Edition)*, 55(1), 51. <https://doi.org/10.2307/1534765>
- Duru-Bellat, M., & Mingat, A. (1988). *Le déroulement de la scolarité au collège: le contexte " fait des différences "...* (Vol. 29, Numéro 4). <https://about.jstor.org/terms>

- Duru-Bellat, M., & Mingat, A. (1997). La constitution de classes de niveau dans les collèges; les effets pervers d'une pratique à visée égalisatrice. *Revue Française de Sociologie*, 38(4), 759. <https://doi.org/10.2307/3322627>
- Duthoit, E., Mailles Viard Metz, S., & Pelissier, C. (2012). Processus d'aide en contexte d'apprentissage: une adaptation pour individualiser et personnaliser. *Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation*, 19(1), 389-406. <https://doi.org/10.3406/stice.2012.1051>
- Education nationale. (2018). *Bulletin officiel n°30 du 26 juillet 2018. Programme d'enseignement du cycle des approfondissements (cycle 4)*.
- Efklides, A. (2001). Metacognitive Experiences in Problem Solving. In A. Efklides, J. Kuhl, & R. Sorrentino (Éds.), *Trends and Prospects in Motivation Research* (p. 297-323). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47676-2_16
- Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review*, 1(1), 3-14. <https://doi.org/10.1016/j.edurev.2005.11.001>
- Efklides, A. (2011). Interactions of Metacognition With Motivation and Affect in Self-Regulated Learning: The MASRL Model. *Educational Psychologist*, 46(1), 6-25. <https://doi.org/10.1080/00461520.2011.538645>
- Efklides, A., Schwartz, B. L., & Brown, V. (2017). Motivation and affect in self-regulated learning: does metacognition play a role? In *Handbook of self-regulation of learning and performance* (p. 64-82). Routledge.
- El Youssifi, M., Boumahdi, A., & El Yousfi, B. (2022). Évaluation des effets des biais de la correction sur l'objectivité du jugement évaluatif en enseignement supérieur. *Revue Marocaine de l'Évaluation et de la Recherche Educative*, 8(8), 258-278.
- Eugène, M. (2023). Appréhender les non-lecteurs scolaires : questions de méthode. *Éducation et didactique*, 17, 43-54. <https://doi.org/10.4000/educationdidactique.11121>
- Ewing, L.-A. (2021). Rethinking Higher Education Post COVID-19. In J. Lee & S. H. Han (Éds.), *The Future of Service Post-COVID-19 Pandemic* (Vol. 1, p. 37-54). Springer. https://doi.org/10.1007/978-981-33-4126-5_3
- Falchikov, N., & Boud, D. (1989). Student Self-Assessment in Higher Education: A Meta-Analysis. *Review of Educational Research*, 59(4), 395-430. <https://doi.org/10.3102/00346543059004395>

- Fauré, L., Gardiès, C., & Marcel, J.-F. (2018). L'usage des technologies numériques pour l'évaluation des pratiques d'enseignement-apprentissage. *L'évaluation en éducation et en formation face aux transformations des sociétés contemporaines*, 319-321.
- Fenouillet, F. (2009). Vers une approche intégrative des théories de la motivation. In P. Carré & F. Fenouillet (Éds.), *Traité de psychologie de la motivation* (p. 305-338).
- Flament, C., & Rouquette, M.-L. (2003). *Anatomie des idées ordinaires. Comment étudier les représentations sociales*. Armand Colin.
- Flavell, J. H. (1976). Metacognitive aspects of problem-solving. In L. B. Resnick (Éd.), *The nature of intelligence*. Lawrence Erlbaum associates.
- Flores, M. A., Veiga Simão, A. M., Barros, A., & Pereira, D. (2015). Perceptions of effectiveness, fairness and feedback of assessment methods: a study in higher education. *Studies in Higher Education*, 40(9), 1523-1534. <https://doi.org/10.1080/03075079.2014.881348>
- Forget, A. (2017a). Quels sont les différents types de différenciation pédagogique dans la classe? In Cnesco (Éd.), *Notes remises dans le cadre de la conférence de consensus du Cnesco et de l'Ifé/Ens de Lyon « Différenciation pédagogique : comment adapter l'enseignement pour la réussite de tous les élèves ? »* (p. 17-25). <http://www.cnesco.fr/fr/differenciation-pedagogique/>
- Forget, A. (2017b, mars). La différenciation dans l'enseignement : état des lieux et questionnement. *Conférence de consensus. Différenciation pédagogique: Comment adapter l'enseignement pour la réussite de tous les élèves*. <http://www.cnesco.fr/fr/differenciation-pedagogique/>
- Forget, A., & Lehraus, K. (2015). La différenciation en classe : qu'en est-il des pratiques réelles des enseignants? *Formation et profession*, 23(3), 70. <https://doi.org/10.18162/fp.2015.287>
- Forrer, D. A., & Wyant, N. A. (2015). Improving Writing Through The Peer-To-Peer Evaluation Process. In *Journal of Business & Economics Research-Third Quarter* (Vol. 13, Numéro 3).
- Foster, R. H., Russell, C. C., Dillon, R., Bitsko, M. J., Godder, K., & Stern, M. (2014). Relations Among Optimism, Perceived Health Vulnerability, and Academic, Self-Regulatory, and Social Self-Efficacy in Adolescent Survivors of Childhood Cancer. *Journal of Psychosocial Oncology*, 32(2), 207-223. <https://doi.org/10.1080/07347332.2013.874000>
- Freinet, C. (1980). *Les techniques Freinet de l'école moderne*. Armand Colin.

- Gallego Noche, B., Quesada Serra, V., Gómez Ruiz, M. A., & Cubero Ibáñez, J. (2017). La evaluación y retroalimentación electrónica entre iguales para la autorregulación y el aprendizaje estratégico en la universidad: la percepción del alumnado. *REDU. Revista de Docencia Universitaria*, 15(1), 127. <https://doi.org/10.4995/redu.2017.5991>
- Garcia-Debanc, C. (1993). Enseignement de la langue et production d'écrits. *Pratiques*, 77(1), 3-23. <https://doi.org/10.3406/prati.1993.1681>
- Gardes, M.-L., & Courtier, P. (2018). Quelle manipulation, représentation et communication dans les ateliers Montessori de première numération? *Grand N, Revue de mathématiques, de sciences et technologie pour les maîtres de l'enseignement primaire*, 101, 83-105.
- Gardner, H. J., & Martin, M. A. (2007). Analyzing Ordinal Scales in Studies of Virtual Environments: Likert or Lump It! *Presence: Teleoperators and Virtual Environments*, 16(4), 439-446. <https://doi.org/10.1162/pres.16.4.439>
- Gauthier, C., Bissonnette, S., & Richard, M. (2007). L'enseignement explicite. In V. Dupriez & G. Chapelle (Éds.), *Enseigner* (p. 107-116). Presses universitaires de France. <http://3e.voie.free.fr>
- Gillies, R. M. (2004). The effects of cooperative learning on junior high school students during small group learning. *Learning and Instruction*, 14(2), 197-213. [https://doi.org/10.1016/S0959-4752\(03\)00068-9](https://doi.org/10.1016/S0959-4752(03)00068-9)
- Go, N. (2013). Approche coopérative et complexe en éducation. In M. Sumputh & F. Fourcade (Éds.), *Oser la pédagogie coopérative complexe: de l'école à l'université* (p. 47-80). Chronique Sociale.
- Gombert, J. E. (1990). Le développement métalinguistique. In *Psychologie d'aujourd'hui*. Presses universitaires de France.
- Grangeat, M. (1999). Processus métacognitifs et différenciation pédagogique. In C. Depover & B. Noël (Éds.), *L'évaluation des compétences et des processus cognitifs, modèles, pratiques et contextes* (p. 115-127). De Boeck.
- Grangeat, M. (2016). Lev S. Vygotski : l'apprentissage par le groupe. In M. Fournier (Éd.), *Éduquer et former : connaissances et débats en éducation et formation* (p. 134-141). Éditions Sciences humaines.
- Grech, V., & Calleja, N. (2018). WASP (Write a Scientific Paper): Parametric vs. non-parametric tests. *Early Human Development*, 123, 48-49. <https://doi.org/10.1016/j.earlhumdev.2018.04.014>
- Håkansson Lindqvist, M., Mozelius, P., Jaldemark, J., & Cleveland-Innes, M. (2020). A Literature Review of Higher Education Reform and Lifelong Learning in a Digital Era.

- EDEN Conference Proceedings, 1*, 189-197. <https://doi.org/10.38069/edenconf-2020-rw-0021>
- Hall, G. E., & Hord, S. M. (2006). *Implementing change: Patterns, principles, and potholes*. Pearson/Allyn and Bacon.
- Harpe, S. E. (2015). How to analyze Likert and other rating scale data. *Currents in Pharmacy Teaching and Learning*, 7(6), 836-850. <https://doi.org/10.1016/j.cptl.2015.08.001>
- Harris, J. E., Boushey, C., Bruemmer, B., & Archer, S. L. (2008). Publishing Nutrition Research: A Review of Nonparametric Methods, Part 3. *Journal of the American Dietetic Association*, 108(9), 1488-1496. <https://doi.org/10.1016/j.jada.2008.06.426>
- Hawe, E., & Dixon, H. (2017). Assessment for learning: a catalyst for student self-regulation. *Assessment & Evaluation in Higher Education*, 42(8), 1181-1192. <https://doi.org/10.1080/02602938.2016.1236360>
- Hayes, A. F., & Coutts, J. J. (2020). Use Omega Rather than Cronbach's Alpha for Estimating Reliability. But.... *Communication Methods and Measures*, 14(1), 1-24. <https://doi.org/10.1080/19312458.2020.1718629>
- Henri, F., & Lundgren-Cayrol, K. (2001). *Apprentissage collaboratif à distance : pour comprendre et concevoir les environnements d'apprentissage virtuels*. Presses de l'Université du Québec.
- Hoffman, J. (2002). Flexible Grouping Strategies in the Multiage Classroom. *Theory Into Practice*, 41(1), 47-52. https://doi.org/10.1207/s15430421tip4101_8
- Hortigüela Alcalá, D., Palacios Picos, A., & López Pastor, V. (2019). The impact of formative and shared or co-assessment on the acquisition of transversal competences in higher education. *Assessment & Evaluation in Higher Education*, 44(6), 933-945. <https://doi.org/10.1080/02602938.2018.1530341>
- Houssaye, J. (2012). La gestion pédagogique des différences entre les élèves : variations françaises. *Carrefours de l'éducation*, n° 34(2), 227-245. <https://doi.org/10.3917/cdle.034.0227>
- Ishikawa, Y., Umemoto, T., Tsubota, Y., Fujiwara, S., Suto, A., & Nishiyama, K. (2021). *Achieving Student Engagement in Learning: Utilizing a Rubric-Based Assessment System for Visualizing Learners' Self-, Peer, and Teacher Assessments* (p. 292-300). https://doi.org/10.1007/978-3-030-77889-7_20
- Jacques, M.-P., & Poibeau, T. (2010). Étudier des structures de discours : préoccupations pratiques et méthodologiques. *Corela*, 8-2. <https://doi.org/10.4000/corela.1855>

- Jacques, M.-P., & Rinck, F. (2017). Un « corpus de littéracie avancée : résultat et point de départ. *Corpus*, 16. <https://doi.org/10.4000/corpus.2806>
- Jamieson, S. (2004). Likert scales: how to (ab)use them. *Medical Education*, 38(12), 1217-1218. <https://doi.org/10.1111/j.1365-2929.2004.02012.x>
- Jason, L., & Glenwick, D. (2016). *Handbook of methodological approaches to community-based research: Qualitative, quantitative, and mixed methods*. Oxford university press.
- Joannidès, R. (2014). *L'écriture électronique des collégiens : quelles questions pour la didactique du français ?* Université de Rouen.
- Johnson, D. W., & Johnson, R. T. (2002). Learning Together and Alone: Overview and Meta-analysis. *Asia Pacific Journal of Education*, 22(1), 95-105. <https://doi.org/10.1080/02188790220220110>
- Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher*, 38(5), 365-379. <https://doi.org/10.3102/0013189X09339057>
- Johnson, D. W., & Johnson, R. T. (2017). Cooperative learning. *Congrès international Innovation Education*.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research*, 1(2), 112-133. <https://doi.org/10.1177/1558689806298224>
- Juita, H. R., & Widiyanto, S. (2019). The Effectiveness of Cooperative Learning Methods: A case study of writing learning at Junior High School. *Proceedings of the Second Conference on Language, Literature, Education, and Culture (ICOLLITE 2018)*. <https://doi.org/10.2991/icollite-18.2019.58>
- Kahn, S. (2010). Les réformes contre le redoublement ou les limites des recherches non collaboratives. *Recherches en éducation, HSI*. <https://doi.org/10.4000/ree.8733>
- Kahn, S. (2015). Différenciation et traitement scolaire des différences. *La nouvelle revue de l'adaptation et de la scolarisation*, 70-71(2), 39. <https://doi.org/10.3917/nras.070.0039>
- Kaplan, J. (2019). Coévaluation entre pairs apprenants comme échafaudage de la régulation du processus de l'apprentissage. *e-JIREF*, 5(1), 47-65.
- Kayacan, A., & Razi, S. (2017). Digital self-review and anonymous peer feedback in Turkish high school EFL writing. *Journal of Language and Linguistic Studies*, 13(2), 561-577.
- Kerzil, J. (2009). Constructivisme. In *L'ABC de la VAE* (p. 112-113). Érès.

- Koehler, M. J., & Mishra, P. (2005). What happens when teachers design educational technology? The development of technological pedagogical content knowledge. *The Journal of Educational Computing Research*, 32(2), 131-152.
- Koehler, M. J., & Mishra, P. (2009). What is technological pedagogical content knowledge? Contemporary Issues in Technology and Teacher Education. *Contemporary Issues in Technology and Teacher Education*, 9(1), 60-70. <http://www.tpck.org/>.
- Kuhl, J. (1987). Action Control: The Maintenance of Motivational States. In *Motivation, Intention, and Volition* (p. 279-291). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-70967-8_19
- Lacelle, N., & Lebrun, M. (2018). La formation à l'écriture numérique : 20 recommandations pour passer du papier à l'écran. *Revue de recherches en littératie médiatique multimodale*, 3. <https://doi.org/10.7202/1047131ar>
- Laffitte, R., & Groupe Vers la pédagogie institutionnelle. (2006). *Essais de pédagogie institutionnelle: l'école, un lieu de recours possible pour l'enfant et ses parents*. Champ social éditions.
- Lafontaine, D. (2017). Quels sont les différents types de différenciation structurelle dans les écoles ou les établissements scolaires? Que sait-on de leurs effets? In Cnesco (Éd.), *Notes remises dans le cadre de la conférence de consensus du Cnesco et de l'Ifé/Ens de Lyon « Différenciation pédagogique : comment adapter l'enseignement pour la réussite de tous les élèves ? »* (p. 9-16).
- Laveault, D., & Allal, L. (2016). Implementing Assessment for Learning: Theoretical and Practical Issues. In *Assessment for Learning: Meeting the Challenge of Implementation* (p. 1-18). Springer Cham. https://doi.org/10.1007/978-3-319-39211-0_1
- Le Boedec, K. (2016). Sensitivity and specificity of normality tests and consequences on reference interval accuracy at small sample size: a computer-simulation study. *Veterinary Clinical Pathology*, 45(4), 648-656. <https://doi.org/10.1111/vcp.12390>
- Lecomte, J. (2004). Les applications du sentiment d'efficacité personnelle. *Savoirs, Hors série*(5), 59-90. <https://doi.org/10.3917/savo.hs01.0059>
- Legrand, L. (1995). *Les différenciations de la pédagogie*. Presses universitaires de France.
- Lescouarch, L. (2010). Spécificité actuelle d'une approche alternative : la pédagogie Freinet. *Spirale. Revue de recherches en éducation*, 45(1), 81-101. <https://doi.org/10.3406/spira.2010.1159>
- Lescouarch, L. (2018). *Construire des situations pour apprendre*. ESF. <https://doi.org/10.14375/NP.9782710137511>

- Lescouarch, L. (2020). Démarche collaborative entre praticiens et chercheurs en recherche pédagogique : l'exemple de la structuration d'une filière de pédagogie coopérative dans un collège en Normandie. *Questions vives recherches en éducation*, N° 33. <https://doi.org/10.4000/questionsvives.4911>
- Lescouarch, L. (2021). Classes coopératives multi-âges : une approche alternative des différenciations pédagogiques. *Éducation et socialisation*, 59. <https://doi.org/10.4000/edso.13738>
- Leslie, K. (2008). Alcohol and drug use among teenagers. *Canadian Medical Association Journal*, 178(2), 149-149. <https://doi.org/10.1503/cmaj.071622>
- Lignier, B., Petot, J. M., Plaisant, O., & Zebdi, R. (2016). Validation du Big Five Inventory français (BFI-Fr) dans un contexte d'hétéro-évaluation. *Annales Medico-Psychologiques*, 174(6), 436-441. <https://doi.org/10.1016/j.amp.2015.07.025>
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.
- Llera, E., Zabalza, I., Scarpellini, S., Andrés, J. A., & Ortega, R. (2015). Analysis of the co-evaluation technique based on collaborative rubric as a learning tool. *7th International Conference on Education and New Learning Technologies*, 3305-3314.
- Loisy, C., Mailles-Viard – Metz, S., & Breton, H. (2011). Se connaître et s'orienter grâce au e-portfolio. *Revue internationale des technologies en pédagogie universitaire*, 8(1-2), 69. <https://doi.org/10.7202/1005785ar>
- Longhurst, N., & Norton, L. S. (1997). Self-assessment in coursework essays. *Studies in Educational Evaluation*, 23(4), 319-330. [https://doi.org/10.1016/S0191-491X\(97\)86213-X](https://doi.org/10.1016/S0191-491X(97)86213-X)
- Lu, J., & Law, N. (2012). Online peer assessment: effects of cognitive and affective feedback. *Instructional Science*, 40(2), 257-275. <https://doi.org/10.1007/s11251-011-9177-2>
- Lüftenegger, M., Schober, B., van de Schoot, R., Wagner, P., Finsterwald, M., & Spiel, C. (2012). Lifelong learning as a goal – Do autonomy and self-regulation in school result in well prepared pupils? *Learning and Instruction*, 22(1), 27-36. <https://doi.org/10.1016/j.learninstruc.2011.06.001>
- Lynd, R. S., & Lynd, H. M. (1959). *Middletown: A Study in Cultural Conflicts*. Harcourt, Brace.
- Mailles Viard Metz. (2015a). *Autonomie et apprentissage universitaire: Aides et outils*. Iste Editions.

- Mailles Viard Metz, S. (2015b). L'aide du numérique aux activités d'auto-évaluation. *Revue internationale de pédagogie de l'enseignement supérieur*, 31(3).
<https://doi.org/10.4000/ripes.991>
- Mailles Viard Metz, S., Vayre, É., & Pelissier, C. (2014). Scénario pédagogique pour la réalisation d'une maquette d'EPA par des étudiants de 1re année de Licence : une aide à l'autorégulation de l'apprentissage ? *Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation*, 21(1), 391-418.
<https://doi.org/10.3406/stice.2014.1104>
- Maity, S., Sahu, T. N., & Sen, N. (2021). Panoramic view of digital education in COVID-19: A new explored avenue. *Review of Education*, 9(2), 405-423.
<https://doi.org/10.1002/rev3.3250>
- Mangenot, F. (2000). Contexte et conditions pour une réelle production d'écrits en ALAO. *Alsic*, Vol. 3, n° 2, 187-206. <https://doi.org/10.4000/alsic.1833>
- Mangenot, F. (2012). Écrire avec l'ordinateur: du traitement de texte au web social. *Français dans le monde. Recherches et applications*, 51, 107-117.
- Marcel, J.-F. (2010). Des tensions entre le « sur » et le « pour » dans la recherche en éducation : question(s) de posture(s). *Éducation et socialisation*, 27-28, 41-64.
<https://doi.org/10.4000/edso.16184>
- Marcel, J.-F. (2012). Contribution à une ingénierie de la commande. *TransFormations-Recherches en Education et Formation des Adultes*, 8.
- Marcel, J.-F. (2016a). *La recherche-intervention : question(s) de forme(s)* (p. 223-230).
<https://doi.org/10.3917/edagri.marce.2016.01.0223>
- Marcel, J.-F. (2016b). *Recherche en sciences de l'éducation et demande sociale* (p. 23-33).
<https://doi.org/10.3917/edagri.marce.2016.01.0023>
- Marchand, P., & Ratinaud, P. (2012). L'analyse de similitude appliquée aux corpus textuels: les primaires socialistes pour l'élection présidentielle française (septembre-octobre 2011). *Actes des 11eme Journées internationales d'Analyse statistique des Données Textuelles. JADT, 2012*, 687-699.
- Marín, V. I., & Pérez Garcias, A. (2016). Collaborative e-Assessment as a Strategy for Scaffolding Self-Regulated Learning in Higher Education. In *Formative Assessment, Learning Data Analytics and Gamification* (p. 3-24). Elsevier.
<https://doi.org/10.1016/B978-0-12-803637-2.00001-4>

- Marquet, P. (2003). *L'impact des TIC dans l'enseignement et la formation : Mesures, modèles et méthodes. Contribution à l'évolution du paradigme comparatif des usages de l'informatique en pédagogie*. <https://tel.archives-ouvertes.fr/edutice-00000384>
- Marquet, P. (2005a). Intérêt du concept de conflit instrumental pour la compréhension des usages des EIAH. *EIAH 2005 (Environnements Informatiques pour l'Apprentissage Humain)*, 383-388.
- Marquet, P. (2005b). Lorsque le développement des TIC et l'évolution des théories de l'apprentissage se croisent. *Savoirs*, 9(3), 105. <https://doi.org/10.3917/savo.009.0105>
- Marquet, P. (2011). e-Learning et conflit instrumental. *Recherche & formation*, 68, 31-46. <https://doi.org/10.4000/rechercheformation.1499>
- Marquet, P., & Coulibaly, B. (2011). The concept of instrumental conflict: an application of the theory of activity to computer-supported teaching-learning situations. In A. Méndez-Vilas (Éd.), *Education in a technological world : communicating current and emerging research and technological efforts* (p. 478-490). Formatex. <https://hal.archives-ouvertes.fr/hal-03224761/>
- Masson, J., & Fenouillet, F. (2013). Relation entre sentiment d'efficacité personnelle et résultats scolaires à l'école primaire : Construction et validation d'une échelle. *Enfance*, N° 4(4), 374-392. <https://doi.org/10.3917/enf1.134.0374>
- McDonald, R. P. (2013). *Test Theory*. Psychology Press. <https://doi.org/10.4324/9781410601087>
- McDonald, R. P. (2014). *Factor Analysis and Related Methods*. Psychology Press. <https://doi.org/10.4324/9781315802510>
- McIntosh, M. E. (1997). Formative Assessment in Mathematics. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 71(2), 92-96. <https://doi.org/10.1080/00098659709599333>
- Meirieu, P. (2015). *L'école, mode d'emploi: des méthodes actives à la pédagogie différenciée*. ESF Sciences Humaines.
- Meirieu, P., & le Bars, S. (2001). *La machine-école*. Gallimard.
- Meirieu, P., & Wagnon, S. (2018). Pédagogie : la fin de la naïveté ! *Tréma*, 50. <https://doi.org/10.4000/trema.4227>
- Merrett, F. (2006). Reflections on the Hawthorne Effect. *Educational Psychology*, 26(1), 143-146. <https://doi.org/10.1080/01443410500341080>
- Meusen-Beekman, K. D., Joosten-ten Brinke, D., & Boshuizen, H. P. A. (2016). Effects of formative assessments to develop self-regulation among sixth grade students: Results

- from a randomized controlled intervention. *Studies in Educational Evaluation*, 51, 126-136. <https://doi.org/10.1016/j.stueduc.2016.10.008>
- Millard, D., Newman, D., & Sinclair, P. (2008). PeerPigeon: A web application to support generalised peer review. *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, 3824-3836.
- Miller-Cochran, S., & Gierdowski, D. (2013). Making Peace with the Rising Costs of Writing Technologies: Flexible Classroom Design as a Sustainable Solution. *Computers and Composition*, 30(1), 50-60. <https://doi.org/10.1016/j.compcom.2012.12.002>
- Priorités du plan national de formation 2017-2018 du ministère de l'éducation nationale. Bulletin officiel n°27, (2017).
- Moinard, P. (2017). Commenter des textes littéraires en collaboration sur des forums et des blogs dans le secondaire. *Le français aujourd'hui*, N° 196(1), 71-80. <https://doi.org/10.3917/lfa.196.0071>
- Monseur, C., & Baye, A. (2016). *Quels apports des données PISA pour l'analyse des inégalités scolaires*. www.cnesco.fr
- Morin, E. (1999). *La tête bien faite. Repenser la réforme. Réformer la pensée*. Editions du Seuil.
- Mostert, M., & Snowball, J. D. (2013). Where angels fear to tread: online peer-assessment in a large first-year class. *Assessment & Evaluation in Higher Education*, 38(6), 674-686. <https://doi.org/10.1080/02602938.2012.683770>
- Mottier Lopez, L., & Allal, L. (2008). Le jugement professionnel en évaluation: un acte cognitif et une pratique sociale située. *Schweizerische Zeitschrift für Bildungswissenschaften*, 30(3), 465-482.
- Mowl, G., & Pain, R. (1995). Using Self and Peer Assessment to Improve Students' Essay Writing: a Case Study from Geography. *Innovations in Education and Training International*, 32(4), 324-335. <https://doi.org/10.1080/1355800950320404>
- Müller, C., & Mildemberger, T. (2021). Facilitating flexible learning by replacing classroom time with an online learning environment: A systematic review of blended learning in higher education. *Educational Research Review*, 34, 100394. <https://doi.org/10.1016/j.edurev.2021.100394>
- Murillo, F. J., & Hidalgo, N. (2015). Enfoques fundamentantes de la evaluación de estudiantes para la justicia social. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 43-61.
- Navarro, D., & Foxcroft, D. (2022). *Learning statistics with jamovi: a tutorial for psychology students and other beginners*. <http://learnstatswithjamovi.com>

- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. <https://doi.org/10.1080/03075070600572090>
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, 15(5), 625-632. <https://doi.org/10.1007/s10459-010-9222-y>
- Nunnally, J., & Bernstein, I. (1994). *Psychometric theory* (3rd éd.). McGraw-Hill.
- Oberti, M. (2007). *L'école dans la ville*. Presses de Sciences Po. <https://doi.org/10.3917/scpo.obert.2007.01>
- O'Brien, S. F., Osmond, L., & Yi, Q.-L. (2015). How do I interpret a p value? *Transfusion*, 55(12), 2778-2782. <https://doi.org/10.1111/trf.13383>
- OCDE. (2014). *Résultats de TALIS 2013 : Une perspective internationale sur l'enseignement et l'apprentissage*. Éditions OCDE. <https://doi.org/10.1787/9789264214293-fr>
- OCDE. (2019a). *PISA 2018 Results (Volume III)*. OCDE. <https://doi.org/10.1787/acd78851-en>
- OCDE. (2019b). *Programme international pour le suivi des acquis des élèves (PISA), résultats du PISA 2018 pour la France*.
- Orsmond, P., Maw, S. J., Park, J. R., Gomez, S., & Crook, A. C. (2013). Moving feedback forward: theory to practice. *Assessment & Evaluation in Higher Education*, 38(2), 240-252. <https://doi.org/10.1080/02602938.2011.625472>
- Oubrayrie-Roussel, N., & Safont-Mottay, C. (2001). Conduites à risques et dévalorisation de soi: Etude de la consommation de toxiques (tabac, alcool, drogue) chez les adolescents scolarisés. *Psicologia, Saúde & Doenças*, 1(2), 59-75.
- Ouherrou, N. (2022). *Jeux Sérieux et Acceptabilité : Vers une Approche de Conception Participative en Éducation Inclusive*. Université Paul Valéry-Montpellier III.
- Pajares, F., Hartley, J., & Valiante, G. (2001). Response Format in Writing Self-Efficacy Assessment: Greater Discrimination Increases Prediction. *Measurement and Evaluation in Counseling and Development*, 33(4), 214-221. <https://doi.org/10.1080/07481756.2001.12069012>
- Panadero, E., Jonsson, A., & Strijbos, J.-W. (2016). *Scaffolding Self-Regulated Learning Through Self-Assessment and Peer Assessment: Guidelines for Classroom Implementation* (p. 311-326). https://doi.org/10.1007/978-3-319-39211-0_18
- Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, 8(3), 293-316. [https://doi.org/10.1016/0361-476X\(83\)90018-8](https://doi.org/10.1016/0361-476X(83)90018-8)

- Pastorelli, C., Caprara, G. V., Barbaranelli, C., Rola, J., Rozsa, S., & Bandura, A. (2001). The structure of children's perceived self-efficacy: A cross-national study. *European Journal of Psychological Assessment, 17*(2), 87-97.
- Penloup, M.-C. (2017a). Didactique de l'écriture : le déjà-là des pratiques d'écriture numérique. *Le français aujourd'hui, N° 196*(1), 57-70. <https://doi.org/10.3917/lfa.196.0057>
- Penloup, M.-C. (2017b). Didactique de l'écriture : le déjà-là des pratiques d'écriture numérique. *Le français aujourd'hui, N° 196*(1), 57-70. <https://doi.org/10.3917/lfa.196.0057>
- Pentecouteau, H. (2020). Comment chercher ensemble ? De l'utilité sociale des sciences humaines et sociales. *Education Permanente, 225*(2020-4).
- Perrenoud, P. (1998). *L'évaluation des élèves. De la fabrication de l'excellence à la régulation des apprentissages*. De Boeck.
- Perrenoud, P. (2001). Évaluation formative et évaluation certificative : postures contradictoires ou complémentaires ? *Formation professionnelle suisse, 4*, 25-28.
- Perrenoud, P. (2014). *Pédagogie différenciée*. ESF. <https://doi.org/10.14375/NP.9782710126713>
- Pett, M. A. (2015). *Nonparametric statistics for health care research: Statistics for small samples and unusual distributions*. Sage Publications.
- Piaget, J. (1975). *L'équilibration des structures cognitives : problème central du développement*. Presses universitaires de France.
- Pintrich, P. R. (2000). The Role of Goal Orientation in Self-Regulated Learning. In *Handbook of Self-Regulation* (p. 451-502). Elsevier. <https://doi.org/10.1016/B978-012109890-2/50043-3>
- Pintrich, P. R., Wolters, C. A., & Baxter, G. P. (2000). Assessing metacognition and self-regulated learning. In G. Schraw & J. Impara (Éds.), *Issues in the measurement of metacognition* (p. 43-97). Buros Institute.
- Piolat, A., & Roussey, J.-Y. (1995). Le traitement de texte : un environnement d'apprentissage encore à expérimenter. *Repères, 11*(1), 87-102. <https://doi.org/10.3406/reper.1995.2146>
- Plane, S. (1996). Écriture, réécriture et traitement de texte. In J. David & S. Plane (Éds.), *L'apprentissage de l'écriture de l'école au collège* (p. 37-78). Presses Universitaires de France. <https://doi.org/10.3917/puf.plane.1996.01.0037>
- Poisson, D., Moisan, A., & Carré, P. (2015). *L'autoformation : perspectives de recherche*. Presses Universitaires de France.

- Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: applications to practice* (Vol. 892). Pearson/Prentice Hall.
- Prost, A. (2001). La démocratisation de l'enseignement en France depuis la Seconde Guerre mondiale. *Swiss Journal of Educational Research*, 23(1), 73-94. <https://doi.org/10.24452/sjer.23.1.4597>
- Pryor, J., & Torrance, H. (1997). Formative Assessment in the Classroom: Where Psychological Theory Meets Social Practice. *Social Psychology of Education*, 2(2), 151-176. <https://doi.org/10.1023/A:1009654524888>
- Quesada, V., Gómez Ruiz, M. Á., Gallego Noche, M. B., & Cubero-Ibáñez, J. (2019). Should I use co-assessment in higher education? Pros and cons from teachers and students' perspectives. *Assessment & Evaluation in Higher Education*, 44(7), 987-1002. <https://doi.org/10.1080/02602938.2018.1531970>
- Rabardel, P. (1995). *Les hommes et les technologies : une approche cognitive des instruments contemporains*. Armand Colin.
- Ratinaud, P. (2009). IRAMUTEQ: Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires. In *Téléchargeable à l'adresse: <http://www.iramuteq.org>*. <http://www.iramuteq.org>
- Régnard, D. (2012). Utiliser les réseaux sociaux en cours de littérature et de latin. *Le français aujourd'hui*, n°178(3), 99-106. <https://doi.org/10.3917/lfa.178.0099>
- Loi n°89-486 du 10 juillet 1989 d'orientation sur l'éducation, (1989).
- Code de l'Éducation, (2000).
- Reverdy, C. (2016). La coopération entre élèves: des recherches aux pratiques. *Dossier de veille de l'IFÉ*, 114, 1-32.
- Reynaud, L. (2022). *Faire collectif pour apprendre: Des clés pour mettre la coopération au service des apprentissages*. ESF Sciences Humaines.
- Rhéaume, J. (2009). La sociologie clinique comme pratique de recherche en institution. Le cas d'un centre de santé et services sociaux. *Sociologie et sociétés*, 41(1), 195-215. <https://doi.org/10.7202/037913ar>
- Roethlisberger, F. J., & Dickson, W. J. (2003). *Management and the Worker* (Vol. 5). Psychology press.
- Romainville, M., Noël, B., & Wolfs, J.-L. (1995). La métacognition : facettes et pertinence du concept en éducation. *Revue française de pédagogie*, 112(1), 47-56. <https://doi.org/10.3406/rfp.1995.1225>

- Rutkowski, D., & Engel, L. C. (2010). Soft Power and Hard Measures: Large-Scale Assessment, Citizenship and the European Union. *European Educational Research Journal*, 9(3), 381-395. <https://doi.org/10.2304/eerj.2010.9.3.381>
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25(1), 54-67. <https://doi.org/10.1006/ceps.1999.1020>
- Ryan, R. M., & Deci, E. L. (2002). Overview of self-determination theory: An organismic dialectical perspective. In *Handbook of self-determination research* (p. 3-33).
- Rytivaara, A. (2011). Flexible Grouping as a Means for Classroom Management in a Heterogeneous Classroom. *European Educational Research Journal*, 10(1), 118-128. <https://doi.org/10.2304/eerj.2011.10.1.118>
- Samaniego Cho, J. A., Mailles-Viard Metz, S., Vidal, J., & Blanc, N. (2021). Confiance et apprentissage médiatisé par ordinateur: une question d'environnement. *Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation*, 28(2), 117-151.
- Sanojca, E. (2018). *Les compétences collaboratives et leur développement en formation d'adultes : le cas d'une formation hybride*. <https://tel.archives-ouvertes.fr/tel-01709910>
- Schunk, D. H. (1982). Verbal self-regulation as a facilitator of childrens achievement and self-efficacy. *Human learning*, 1(4), 265-277.
- Schunk, D. H., & Ertmer, P. A. (2000). Self-Regulation and Academic Learning. In *Handbook of Self-Regulation* (p. 631-649). Elsevier. <https://doi.org/10.1016/B978-012109890-2/50048-2>
- Schwartz, B. (1973). L'éducation demain. In *Recherches*. Aubier-Montaigne.
- Schwarz, B. (2020). Argumentation et rationalité dans l'étude de textes talmudiques : les interactions dyadiques dans les Yeshivas. *Argumentation et analyse du discours*, 25. <https://doi.org/10.4000/aad.4647>
- Segers, M., Dochy, F., & Cascallar, E. (2003a). *Optimising New Modes of Assessment: In Search of Qualities and Standards* (M. Segers, F. Dochy, & E. Cascallar, Édés.; Vol. 1). Kluwer Academic Publishers. <https://doi.org/10.1007/0-306-48125-1>
- Segers, M., Dochy, F., & Cascallar, E. (2003b). The Era of Assessment Engineering: Changing Perspectives on Teaching and Learning and the Role of New Modes of Assessment. In *Optimising New Modes of Assessment: In Search of Qualities and Standards* (p. 1-12). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-48125-1_1
- Sénèque. (Ier siècle). *Letters on ethics: to Lucilius*. University of Chicago Press.

- Shapiro, S., & Wilk, M. (1965). An analysis of variance test for normality. *Biometrika*, 52(3), 591-611.
- Sharan, Y. (2010). Cooperative Learning for Academic and Social Gains: valued pedagogy, problematic practice. *European Journal of Education*, 45(2), 300-313. <https://doi.org/10.1111/j.1465-3435.2010.01430.x>
- Sheldon, M. R., Fillyaw, M. J., & Thompson, W. D. (1996). The use and interpretation of the Friedman test in the analysis of ordinal-scale data in repeated measures designs. *Physiotherapy Research International*, 1(4), 221-228. <https://doi.org/10.1002/pri.66>
- Sheskin, D. J. (2003). *Handbook of Parametric and Nonparametric Statistical Procedures*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781420036268>
- Singh, P. (2013). P value, statistical significance and clinical significance. *Journal of Clinical and Preventive Cardiology*, 2(4), 202-204.
- Sivan, A. (2000). The Implementation of Peer Assessment: An action research approach. *Assessment in Education: Principles, Policy & Practice*, 7(2), 193-213. <https://doi.org/10.1080/713613328>
- Souza, R. R. de, Toebe, M., Mello, A. C., & Bittencourt, K. C. (2023). Sample size and Shapiro-Wilk test: An analysis for soybean grain yield. *European Journal of Agronomy*, 142, 126666. <https://doi.org/10.1016/j.eja.2022.126666>
- St. Laurent, R., & Turk, P. (2013). The Effects of Misconceptions on the Properties of Friedman's Test. *Communications in Statistics - Simulation and Computation*, 42(7), 1596-1615. <https://doi.org/10.1080/03610918.2012.671874>
- Stefani, L. A. J. (1994). Peer, self and tutor assessment: Relative reliabilities. *Studies in Higher Education*, 19(1), 69-75. <https://doi.org/10.1080/03075079412331382153>
- Strasma, K. (2009). « Spotlighting »: Peer-Response in Digitally Supported First-Year Writing Courses. *Teaching English in the Two Year College*, 37(2), 153.
- Sullivan, G. M., & Artino Jr, A. R. (2013). Analyzing and Interpreting Data From Likert-Type Scales. *Journal of Graduate Medical Education*, 5(4), 541.
- Taglietti, D., Landri, P., & Grimaldi, E. (2021). The big acceleration in digital education in Italy: The COVID-19 pandemic and the blended-school form. *European Educational Research Journal*, 20(4), 423-441. <https://doi.org/10.1177/147490412111021246>
- Tali, F., & Germier, C. (2021). Co-évaluation et évaluation formative dans un carnet de bord en ligne collaboratif. *TransFormations: Recherche en éducation et formation des adultes*, 2(22), 40-60.

- Taous, T. (2020). Forums littéraires en milieu scolaire : analyse de commentaires numériques et de leur influence sur les écrits d'élèves dans le cadre d'une fanfiction. *Revue de recherches en littérature médiatique multimodale*, 12. <https://doi.org/10.7202/1073687ar>
- Thobois-Jacob, L., & Chevry-Pébayle, E. (2020). Développer la recherche d'information et l'autorégulation collective en classe renversée : l'apport des tests d'autoévaluation. *Revue internationale de pédagogie de l'enseignement supérieur*, 36(1). <https://doi.org/10.4000/ripes.2294>
- Thobois-Jacob, L., & Mohib, N. (2020). S'autoévaluer pour mieux se former à l'entrepreneuriat? Exemple d'une démarche d'autoévaluation mise en œuvre dans un module de sensibilisation à l'entrepreneuriat. *Colloque international «Éduquer à l'esprit d'entreprendre, former à l'entrepreneuriat? Enjeux, questions, transformations»*, Lille, France.
- Thobois-Jacob, L., Mohib, N., Leduc, C., & Gensbittel, L. (2020). De l'autoévaluation pour renouveler l'expérience d'enseignement et d'apprentissage. *31ème congrès de l'AIPU (Association Internationale de Pédagogie Universitaire)*, Laval, Canada.
- Thomson, A. M., & Perry, J. L. (2006). Collaboration Processes: Inside the Black Box. *Public Administration Review*, 66(s1), 20-32. <https://doi.org/10.1111/j.1540-6210.2006.00663.x>
- Thomson, A. M., Perry, J. L., & Miller, T. K. (2008). Linking collaboration processes and outcomes: Foundations for advancing empirical theory. In R. O'Leary & L. Bingham (Éds.), *Big Ideas in Collaborative Public Management* (p. 97-120). Sharpe.
- Topping, K. (1998a). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249-276. <https://doi.org/10.2307/1170598>
- Topping, K. (1998b). *Peer Assessment Between Students in Colleges and Universities* (Vol. 68, Numéro 3).
- Topping, K. (2003). Self and Peer Assessment in School and University: Reliability, Validity and Utility. In *Optimising New Modes of Assessment: In Search of Qualities and Standards* (p. 55-87). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-48125-1_4
- Topping, K. J., & Elliot, A. (2000). Formative Peer Assessment of Academic Writing Between Postgraduate Students. In *Assessment & Evaluation in Higher Education* (Vol. 25, Numéro 2).
- Topping, K. J., Smith, E. F., Swanson, I., & Elliot, A. (2000). Formative Peer Assessment of Academic Writing Between Postgraduate Students. *Assessment & Evaluation in Higher Education*, 25(2), 149-169. <https://doi.org/10.1080/713611428>

- Torterat, F. (2015). Les courriels comme écrits de travail. In PUP (Éd.), *L'Écriture experte en questions*. PUP / Sherbrooke Université. <https://www.researchgate.net/publication/299595189>
- Torterat, F., Azaoui, B., Boutevin, C., & Dupuy, C. (2020). Une recherche collaborative en Réseau d'Éducation Prioritaire (REP), entre initiatives de terrain et reconfiguration du travail. *Recherches & éducations*. <https://doi.org/10.4000/rechercheseducations.7984>
- Torterat, F., Blazin, D., Cros, A., Jarrige, I., & Morel, F. (2019). Apports et mise en pratique de démarches coopératives en classe Passerelle. *Recherches. Revue de didactique et de pédagogie du français*, 70.
- Toullec-Théry, M., & Lescouarch, L. (2014). Le maître E en Rased : enjeux, pratiques, perspectives. *La nouvelle revue de l'adaptation et de la scolarisation*, 66(2), 5. <https://doi.org/10.3917/nras.066.0005>
- Trouche, L. (2007). Environnements informatisés d'apprentissage: quelle assistance didactique pour la construction des instruments mathématiques? In *HAL*.
- Tudge, J. R. H., & Winterhoff, P. A. (1993). Vygotsky, Piaget, and Bandura: Perspectives on the Relations between the Social World and Cognitive Development. *Human Development*, 36(2), 61-81. <https://doi.org/10.1159/000277297>
- VandenBos, G. R. (2007). *APA dictionary of psychology*. American Psychological Association.
- Vinck, D. (2009). De l'objet intermédiaire à l'objet-frontière. *Revue d'anthropologie des connaissances*, 3(1). <https://doi.org/10.3917/rac.006.0051>
- Vinutha, H. P., Poornima, B., & Sagar, B. M. (2018). *Detection of Outliers Using Interquartile Range Technique from Intrusion Dataset* (p. 511-518). https://doi.org/10.1007/978-981-10-7563-6_53
- Vygotski, L. (1985). *Pensée et langage*. Messidor.
- Vygotsky, L. S. (1985). *Pensée et langage*. In *Paris: Messidor/Éditions sociales (1re éd. 1934)* (Traduction. F. Sève). Messidor/Éditions sociales.
- Wang, S.-L., & Lin, S. S. J. (2007). The effects of group composition of self-efficacy and collective efficacy on computer-supported collaborative learning. *Computers in Human Behavior*, 23(5), 2256-2268. <https://doi.org/10.1016/j.chb.2006.03.005>
- Warren, S. J., Dondlinger, M. J., & Barab, S. A. (2008). A MUVE Towards PBL Writing. *Journal of Research on Technology in Education*, 41(1), 113-140. <https://doi.org/10.1080/15391523.2008.10782525>
- Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research*, 13(1), 21-39. [https://doi.org/10.1016/0883-0355\(89\)90014-1](https://doi.org/10.1016/0883-0355(89)90014-1)

- Wen, M. L., & Tsai, C.-C. (2006). University Students' Perceptions of and Attitudes Toward (Online) Peer Assessment. *Higher Education*, 51(1), 27-44. <https://doi.org/10.1007/s10734-004-6375-8>
- Xiao, L., & Hau, K.-T. (2023). Performance of Coefficient Alpha and Its Alternatives: Effects of Different Types of Non-Normality. *Educational and Psychological Measurement*, 83(1), 5-27. <https://doi.org/10.1177/00131644221088240>
- Yang, Y.-F., & Tsai, C.-C. (2010). Conceptions of and approaches to learning through online peer assessment. *Learning and Instruction*, 20(1), 72-83. <https://doi.org/10.1016/j.learninstruc.2009.01.003>
- Yarrow, F., & Topping, K. J. (2001). Collaborative writing: The effects of metacognitive prompting and structured peer interaction. *British Journal of Educational Psychology*, 71(2), 261-282. <https://doi.org/10.1348/000709901158514>
- Yebbou, J. (2015). Une lecture des résultats de PISA pour la France dans le contexte des pays de l'OCDE. *Administration & Éducation*, N° 145(1), 31. <https://doi.org/10.3917/admed.145.0031>
- Yin, L. (2015). Investigating Robust Estimation and Forecasting of Volatilities of Futures with Interquartile Range Models. *Journal of Finance and Economics*, 3(2), 1-10. <https://doi.org/10.12735/jfe.v3i2p01>
- Zarifian, P. (2009). *Le travail et la compétence : entre puissance et contrôle*. Presses Universitaires de France. <https://doi.org/10.3917/puf.zarif.2009.01>
- Zeller, A., & Marquet, P. (2021a). Approaching Instrumental Conflicts During an Assessment Activity with an Interactive Learning Environment. *ECEL 2021: 20th European Conference on e-Learning*,.
- Zeller, A., & Marquet, P. (2021b, juin). Interface et EIAH : du territoire apprenant au territoire de l'apprenant. *Colloque Inter-AREF Nancy 2021 "Politiques et territoires en éducation et formation : Enjeux, débats et perspectives"*.
- Zimmerman, B. J. (1995). Self-regulation involves more than metacognition: A social cognitive perspective. *Educational Psychologist*, 30(4), 217-221. https://doi.org/10.1207/s15326985ep3004_8
- Zimmerman, B. J. (2000). Attaining Self-Regulation. In *Handbook of Self-Regulation* (p. 13-39). Elsevier. <https://doi.org/10.1016/B978-012109890-2/50031-7>
- Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice*, 41(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

- Zimmerman, B. J. (2008). Investigating Self-Regulation and Motivation: Historical Background, Methodological Developments, and Future Prospects. *American Educational Research Journal*, 45(1), 166-183. <https://doi.org/10.3102/0002831207312909>
- Zimmerman, D. W., & Zumbo, B. D. (1993). Relative Power of the Wilcoxon Test, the Friedman Test, and Repeated-Measures ANOVA on Ranks. *The Journal of Experimental Education*, 62(1), 75-86. <https://doi.org/10.1080/00220973.1993.9943832>
- Zoller, Z., & Ben-Chaim, D. (1997). Student self-assessment in HOCS science examinations: is it compatible with that of teachers? *7th EARLI conference, Athens, Greece*, 26-30.
- Zuber-Skerritt, O., Wood, L., & Kearney, J. (2020). The transformative potential of action learning in community-based research for social action. *Action Learning: Research and Practice*, 17(1), 34-47. <https://doi.org/10.1080/14767333.2020.1712840>

ANNEXES

1 Résultats détaillés

1.1 Résultats préliminaires relatifs à la validité des données

Comme détaillé *supra*, nous testons plusieurs prédictions subalternes (codées « PS »). Ces analyses préliminaires ont pour but d'évaluer la validité des données recueillies.

1.1.1 Expérience préalable en matière d'auto- et d'inter-évaluation (PS 1)

Pour rappel, à T1 – donc en amont de la première session expérimentale avec le dispositif de coévaluation numérique –, deux items du questionnaire en ligne ciblaient l'expérience préalable des élèves participants en matière d'auto- et d'inter-évaluation. Il s'agissait de questions fermées, *i.e.* échelles de Likert à cinq points (minimum = 1 ; maximum = 5). Les réponses étaient situées autour de 3 sur 5 en moyenne, ce qui correspondait à un rang médian dans l'échelle de Likert à cinq points, *i.e.* aux descriptifs :

- « J'ai rarement pratiqué l'auto-évaluation » ;
- « J'ai rarement pratiqué l'inter-évaluation ».

Ces résultats indiquent que les élèves participants avaient déjà une expérience significative de l'auto- et de l'inter-évaluation, ce qui valide l'hypothèse PS 1. La Figure 92 représente ces résultats sous forme de valeurs moyennes, pour les échantillons A, A', B, C et pour le super-échantillon SE.

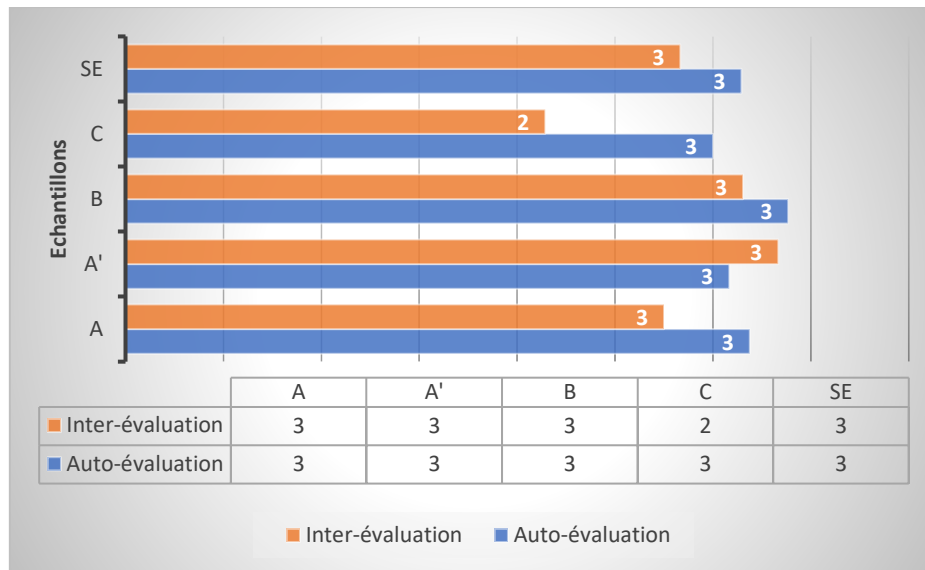


Figure 92 – Résultats des deux items relatifs à l'expérience préalable des élèves participants en matière d'auto- et d'inter-évaluation (valeurs moyennes à T1, échantillons A, A', B, C et SE, N = 62)

1.1.2 Acceptabilité du questionnaire (PS 2)

Pour rappel, à T4, trois items visent à mesurer l'acceptabilité du questionnaire auprès des élèves interrogés, *i.e.* la compréhension, l'apparence et la prise en main du questionnaire. L'analyse descriptive des données pour l'ensemble des participants tous échantillons confondus (N = 57) révèle des moyennes et des médianes élevées (Tableau 43).

Statistiques descriptives			
	Compréhension	Apparence	Prise en main
N	57	57	57
Manquants	0	0	0
Moyenne	4.56	4.35	4.21
Médiane	5	5	4
Ecart-type	0.708	0.916	0.959
Minimum	2	1	1
Maximum	5	5	5

Tableau 43 – Statistiques descriptives pour les 3 questions relatives à l'ergonomie du questionnaire

Pour les trois dimensions évaluées, ces données révèlent ainsi une bonne acceptabilité du questionnaire par les participants, ce qui valide l'hypothèse PS 2.

1.1.3 Compréhension du fonctionnement de l'échelle SEPE (PS 3)

Pour rappel, les 8 items de calibrage du questionnaire SEPE correspondent à des tâches de difficulté croissante, *i.e.* de CAL 1 (tâche la plus facile) à CAL 8 (tâche la plus difficile). La tendance observée dans les réponses des participants et que, plus la difficulté est grande – *i.e.*, ici, plus le poids à soulever est lourd –, plus le sentiment d'efficacité personnelle diminue (Figure 93, Figure 94 et Figure 95). Ces résultats sont convergents à T1 et T3. Ils révèlent une bonne compréhension du principe interprétatif de l'échelle SEPE par les élèves, ce qui valide l'hypothèse PS 3 et suggère que les données collectées via l'échelle SEPE sont fiables.

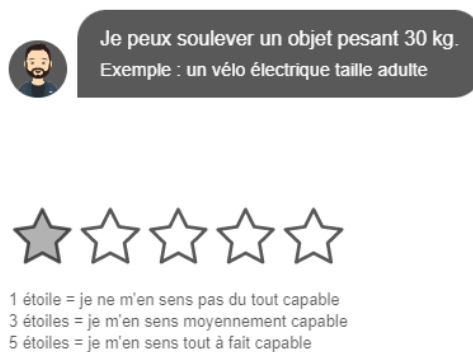


Figure 93 – Exemple de question d'entraînement

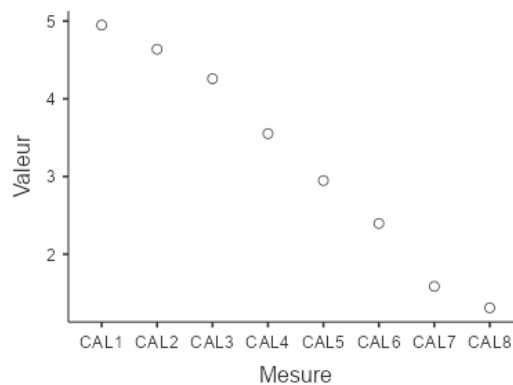


Figure 94 – Résultat du test de calibrage à T1

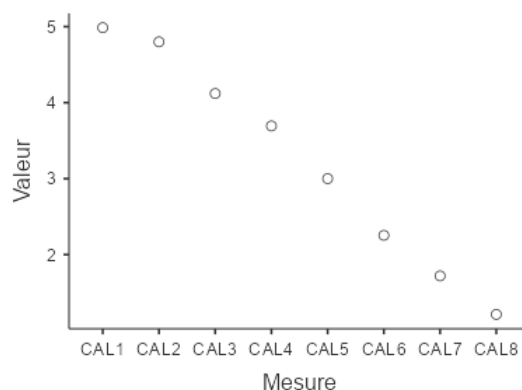


Figure 95 – Résultat du test de calibrage à T3

1.1.4 Cohérence interne de l'échelle SEPE et des sous-échelles afférentes (PS 4)

1.1.4.1 Echelle SEPE

Il est communément admis que la cohérence interne d'une échelle est satisfaisante si les valeurs de l'alpha de Cronbach et de l'oméga de McDonald sont supérieures ou égales à 0.7 (Nunnally & Bernstein, 1994). A ce titre, nos résultats indiquent une cohérence interne très satisfaisante ($\alpha = 0.913$, $\omega = 0.915$). Au-delà de 0.95, on pourrait être en présence d'intercorrélations élevées entre items, avec des effets de redondance nuisant à la qualité de la mesure, ce qui n'est pas le cas ici. Enfin, les résultats convergents obtenus avec les deux coefficients (Tableau 44) valident l'hypothèse PS 4 selon laquelle la cohérence interne de l'échelle SEPE est satisfaisante (Navarro & Foxcroft, 2022, p. 461).

Analyses de cohérence interne		
	α de Cronbach	ω de McDonald
Echelle SEPE	0.913	0.915

Tableau 44 – Tests de fiabilité pour l'échelle SEPE avec SE à T1, T2, T3 et T4 (N = 282)

1.1.4.2 Sous-échelles SEP(scol), SEP(srl) et SEP(coop)

Pour rappel, parmi les sept dimensions incluses dans l'échelle SEPE, trois revêtent un intérêt particulier pour notre étude, en l'occurrence les sous-échelles mesurant :

- l'auto-efficacité perçue pour la réussite scolaire : SEP(scol) ;
- l'auto-efficacité perçue pour l'autorégulation des apprentissages : SEP(srl) ;
- l'auto-efficacité perçue pour les apprentissages coopératifs : SEP(coop).

Pour l'analyse de la cohérence interne, nous avons appliqué à ces trois sous-échelles le même traitement qu'à l'échelle SEPE. Les coefficients obtenus se sont tous révélés satisfaisants au sens de (Nunnally & Bernstein, 1994). Le Tableau 45, le Tableau 46 et le Tableau 47 récapitulent les statistiques de cohérence interne obtenues pour les trois sous-échelles, tandis que la Figure 96 met l'ensemble des résultats en perspective.

Statistiques de fidélité de l'objet		
	α de Cronbach	ω de McDonald
Sous-échelle SEP(scol)	0.787	0.791

Tableau 45 – Tests de cohérence interne pour l'échelle SEP(scol) avec SE à T1, T2, T3 et T4 (N = 282)

Statistiques de fidélité de l'objet		
	α de Cronbach	ω de McDonald
Sous-échelle SEP(srl)	0.856	0.867

Tableau 46 – Tests de cohérence interne pour l'échelle SEP(srl) avec SE à T1, T2, T3 et T4 (N = 282)

Statistiques de fidélité de l'objet		
	α de Cronbach	ω de McDonald
Sous-échelle SEP(coop)	0.703	0.766

Tableau 47 – Tests de cohérence interne pour l'échelle SEP(coop) avec SE à T1, T2, T3 et T4 (N = 282)

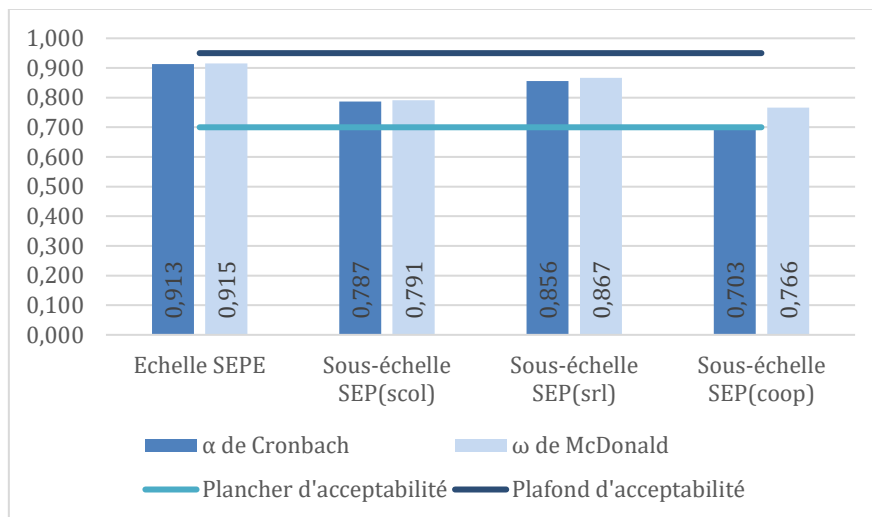


Figure 96 – Cohérence interne pour l'échelle SEPE et les sous-échelles SEP(scol), SEP(srl) et SEP(coop) avec SE à T1, T2, T3 et T4 (N = 282)

1.1.5 Discussion (prédictions subalternes)

Les analyses préliminaires présentées dans ce chapitre nous permettent de valider les quatre prédictions subalternes énoncées *supra*.

En premier lieu, nous avons pu établir que les élèves participants disposaient déjà d'une expérience significative en matière d'auto- et d'inter-évaluation *avant* d'expérimenter la coévaluation dans la présente recherche (PS 1). Ces résultats permettent de rejeter l'hypothèse alternative, selon laquelle les élèves étaient inexpérimentés en la matière, une variable qui aurait biaisé les résultats en suscitant des conflits instrumentaux non directement liés aux artefacts pédagogique (*i.e.* la coévaluation) et numérique (*i.e.* Moodle Atelier) qui font l'objet de la présente recherche.

Nous avons montré ensuite que le questionnaire en ligne était acceptable pour les élèves participants (PS 2). Compte tenu du public pré-adolescent interrogé, ce résultat permet de considérer comme valides les données ainsi recueillies, malgré la relative longueur du questionnaire – *i.e.* entre 40 et 47 items par passation. On peut supposer que la ludification – rendue possible par le logiciel Tripetto®, présenté *supra* – a joué un rôle significatif en ce sens.

En outre, nous avons pu établir que la logique de fonctionnement de l'échelle psychométrique SEPE – et par là même, celle des sous-échelles SEP(scol), SEP(srl) et SEP(coop) – a été bien comprise par les élèves participants (PS 3). Ce résultat est très significatif dans la mesure où il tend à renforcer la validité d'environ 80% de nos données autorapportées – *i.e.* 37 items sur 47. Cela confirme par ailleurs l'importance des questions de calibrage et de leur adaptation au public cible, selon les recommandations de (Bandura, 2006) relativement aux échelles de Likert mesurant le sentiment d'efficacité personnelle.

Enfin, l'approche proposée *supra* pour mesurer la cohérence interne des quatre échelles psychométriques opérationnalisées dans la présente recherche s'est révélée probante. En effet, les résultats convergents obtenus en calculant les coefficients alpha de Cronbach et oméga de McDonald valident l'hypothèse PS 4, selon laquelle les cohérences internes de l'échelle SEPE et des sous-échelles SEP(scol), SEP(srl) et SEP(coop) sont satisfaisantes.

En somme, l'ensemble de ces résultats préliminaires tend à confirmer la validité du protocole de recherche ici mis en œuvre.

1.2 Hypothèse HO 1

1.2.1 Augmentation modérée du score de scripteur de SE sur T2-T4

Statistiques descriptives

	T2	T4
N	58	58
Moyenne	49.3	58.3
Médiane	47.5	58.5
Ecart interquartile	26.3	28.0
Etendue	92	77
Minimum	8	17
Maximum	100	94
W de Shapiro-Wilk	0.978	0.971
Valeur p de Shapiro-Wilk	0.365	0.174

Test t pour échantillons appariés

			Statistique	p		Taille de l'effet
T4	T2	W de Wilcoxon	1163	0.001	Corrélation entre rangs bisériés	0.510

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.2.2 Différence significative entre les scores de scripteurs de l'échantillon C sur T2-T6

Statistiques descriptives

	T2	T4	T6
N	22	22	22
Moyenne	54.3	62.9	71.3
Médiane	56.5	64.0	70.0
Ecart-type	17.3	17.5	19.7
Ecart interquartile	25.3	20.5	34.0
Etendue	60	77	60
Minimum	27	17	40
Maximum	87	94	100
W de Shapiro-Wilk	0.961	0.954	0.925
Valeur p de Shapiro-Wilk	0.511	0.372	0.095

Statistiques descriptives

	T2	T4	T6
Friedman			
χ^2	ddl	p	
8.00	2	0.018	

Comparaison pair-à-pair (Durbin-Conover)

	Statistique		p
T2 - T4	2.00	0.052	
T2 - T6	3.00	0.005	
T4 - T6	1.00	0.323	

1.2.3 Augmentation du SEP(scol) de SE sur T3-T4

Statistiques descriptives

	T3	T4
N	46	46
Moyenne	3.59	3.74
Médiane	3.64	3.86
Ecart interquartile	1.18	1.07
Etendue	3.43	3.14
Minimum	1.57	1.86
Maximum	5.00	5.00
W de Shapiro-Wilk	0.977	0.973
Valeur p de Shapiro-Wilk	0.482	0.353

Test t pour échantillons appariés

			Statistique	p	Taille de l'effet
T4	T3	W de Wilcoxon	412.5	0.001	Corrélation entre rangs bisériés 0.663

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.3 Hypothèse HO 2

1.3.1 Hypothèse HO 2a

1.3.1.1 Augmentation du SEP(srl) de SE sur T1-T2

Statistiques descriptives

	T1	T2
N	46	46
Moyenne	3.64	3.89
Médiane	4.00	4.00
Ecart interquartile	1.38	1.88
Etendue	4.00	4.00
Minimum	1.00	1.00
Maximum	5.00	5.00
W de Shapiro-Wilk	0.899	0.837
Valeur p de Shapiro-Wilk	< .001	< .001

Test t pour échantillons appariés

			Statistique	p		Taille de l'effet
T2	T1	W de Wilcoxon	237.5	0.042	Corrélation entre rangs bisériés	0.462

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.3.1.2 Augmentation du SEP(srl) de SE sur T3-T4

Statistiques descriptives

	T3	T4
N	46	46
Moyenne	3.59	3.70
Médiane	3.60	3.80
Ecart interquartile	0.900	0.900
Etendue	3.90	3.90
Minimum	1.10	1.10
Maximum	5.00	5.00
W de Shapiro-Wilk	0.961	0.958
Valeur p de Shapiro-Wilk	0.126	0.094

Statistiques descriptives

	T3	T4
--	----	----

Test t pour échantillons appariés

			Statistique	p		Taille de l'effet
T4	T3	W de Wilcoxon	487 ^a	0.042	Corrélation entre rangs bisériés	0.384

Note. H_a $\mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.3.1.3 Forte augmentation du SEP(srl) de B sur T1-T2 (scores individuels médians)

Statistiques descriptives

	T1	T2
N	10	10
Moyenne	3.00	3.85
Médiane	3.00	4.25
Ecart interquartile	0.375	1.63
Etendue	3.50	4.00
Minimum	1.00	1.00
Maximum	4.50	5.00
W de Shapiro-Wilk	0.853	0.848
Valeur p de Shapiro-Wilk	0.064	0.054

Test t pour échantillons appariés

			Statistique	p		Taille de l'effet
T2	T1	W de Wilcoxon	40.0	0.041	Corrélation entre rangs bisériés	0.778

Note. H_a $\mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.3.1.4 Forte augmentation du SEP(srl) de B sur T1-T2 (scores individuels moyens)

Statistiques descriptives

	T1	T2
N	10	10
Moyenne	3.25	3.70
Médiane	3.47	4.15
Ecart interquartile	0.425	0.975
Etendue	2.10	2.70

Statistiques descriptives

	T1	T2
Minimum	1.80	1.70
Maximum	3.90	4.40
W de Shapiro-Wilk	0.844	0.788
Valeur p de Shapiro-Wilk	0.049	0.010

Test t pour échantillons appariés

		Statistique	p	Taille de l'effet		
T2	T1	W de Wilcoxon	48.0	0.037	Corrélation entre rangs bisériés	0.745

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.3.2 Hypothèse HO 2b

1.3.2.1 Corrélation entre scores de scripteurs et de coévaluateurs de A à T4

Statistiques descriptives

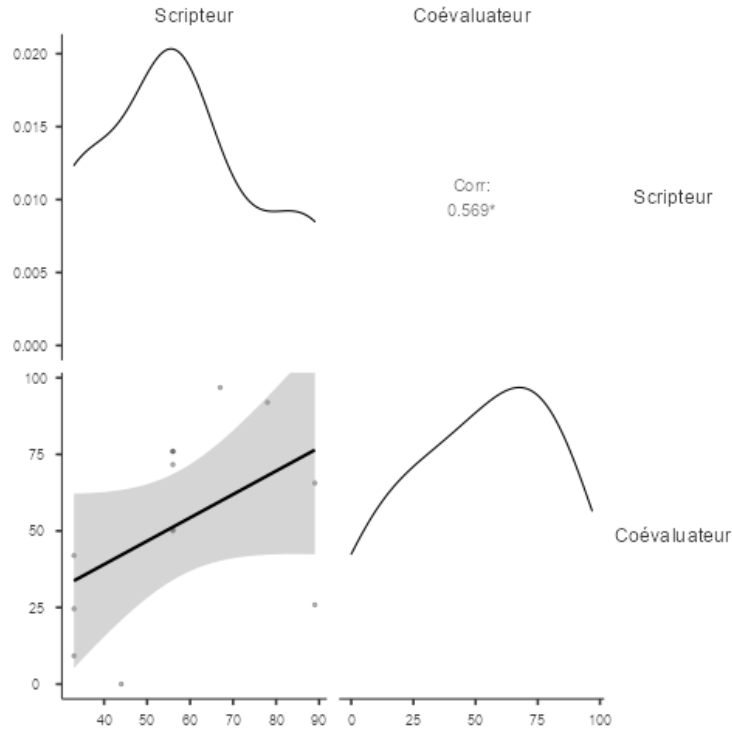
	Scripteur	Coévaluateur
N	13	13
Moyenne	57.4	52.3
Médiane	56	50.5
Ecart interquartile	23.0	50.0
Etendue	56	96.8
Minimum	33	0.00
Maximum	89	96.8
W de Shapiro-Wilk	0.895	0.957
Valeur p de Shapiro-Wilk	0.116	0.709

Matrice de corrélation

		Scripteur	Coévaluateur
Scripteur	Rho de Spearman	—	
	valeur p	—	
	N	—	
Coévaluateur	Rho de Spearman	0.569 *	—
	valeur p	0.042	—
	N	13	—

	Scripteur	Coévaluateur
--	-----------	--------------

Note. * p < .05, ** p < .01, *** p < .001

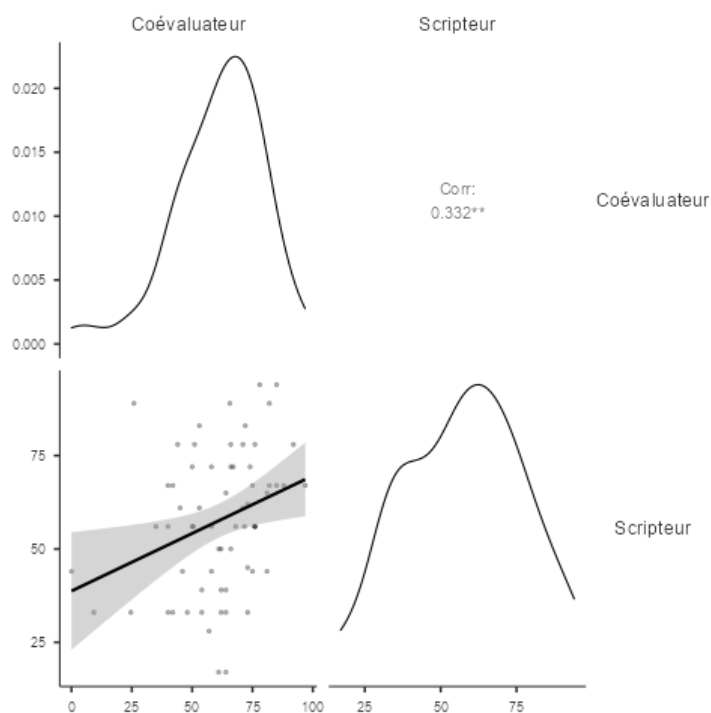


1.3.2.2 Corrélation entre scores de scripteurs et de coévaluateurs de SE à T4

Matrice de corrélation

		Coévaluateur	Scripteur
Coévaluateur	Rho de Spearman	—	
	valeur p	—	
	N	—	
Scripteur	Rho de Spearman	0.332 **	—
	valeur p	0.007	—
	N	65	—

Note. * p < .05, ** p < .01, *** p < .001



1.3.3 Hypothèse HO 2c

1.3.3.1 Augmentation modérée du SEP de SE sur T1-T2

Statistiques descriptives

	T1	T2
N	46	46
Moyenne	3.85	3.95
Médiane	3.95	4.01
Ecart interquartile	0.363	0.473
Etendue	2.49	2.68
Minimum	2.35	2.32
Maximum	4.84	5.00
W de Shapiro-Wilk	0.911	0.924
Valeur p de Shapiro-Wilk	0.002	0.005

Test t pour échantillons appariés

			Statistique	p	Taille de l'effet
T2	T1	W de Wilcoxon	609	0.021	Corrélation entre rangs bisériés 0.413

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.3.3.2 Augmentation modérée du SEP de SE sur T3-T4

Statistiques descriptives

	T3	T4
N	46	46
Moyenne	3.84	3.94
Médiane	3.86	4.01
Ecart interquartile	0.791	0.730
Etendue	3.11	3.14
Minimum	1.89	1.81
Maximum	5.00	4.95
W de Shapiro-Wilk	0.962	0.949
Valeur p de Shapiro-Wilk	0.132	0.044

Test t pour échantillons appariés

			Statistique	p		Taille de l'effet
T4	T3	W de Wilcoxon	798	< .001	Corrélation entre rangs bisériés	0.611

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.3.3.3 Augmentation modérée du SEP de A' sur T3-T4

Statistiques descriptives

	T3	T4
N	12	12
Moyenne	3.81	3.88
Médiane	4.00	4.14
Ecart interquartile	0.426	0.405
Etendue	2.65	2.95
Minimum	1.89	1.81
Maximum	4.54	4.76
W de Shapiro-Wilk	0.784	0.800
Valeur p de Shapiro-Wilk	0.006	0.009

Test t pour échantillons appariés

			Statistique	p		Taille de l'effet
--	--	--	--------------------	----------	--	--------------------------

Statistiques descriptives

			T3	T4		
T4	T3	W de Wilcoxon	64.5	0.049	Corrélation entre rangs bisériés	0.654

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.3.3.4 Augmentation modérée du SEP de C sur T3-T4

Statistiques descriptives

	T3	T4
N	16	16
Moyenne	3.91	4.02
Médiane	3.85	4.04
Ecart interquartile	0.980	1.03
Etendue	2.32	2.27
Minimum	2.68	2.62
Maximum	5.00	4.89
W de Shapiro-Wilk	0.973	0.950
Valeur p de Shapiro-Wilk	0.891	0.489

Test t pour échantillons appariés

			Statistique	p		Taille de l'effet
T4	T3	W de Wilcoxon	108	0.041	Corrélation entre rangs bisériés	0.588

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.3.3.5 Forte augmentation du SEP de B sur T1-T2

Statistiques descriptives

	T1	T2
N	10	10
Moyenne	3.76	4.09
Médiane	3.82	4.15
Ecart interquartile	0.228	0.541
Etendue	1.32	1.05
Minimum	2.78	3.51
Maximum	4.11	4.57
W de Shapiro-Wilk	0.749	0.956

Statistiques descriptives

	T1	T2
Valeur p de Shapiro-Wilk	0.004	0.742

Test t pour échantillons appariés

			Statistique	p	Taille de l'effet
T2	T1	W de Wilcoxon	52.0	0.010	Corrélation entre rangs bisériés
					0.891

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.4 Hypothèse HO 3

1.4.1 SEP(coop)

1.4.1.1 Forte augmentation du SEP(coop) de A' sur T1-T2

Statistiques descriptives

	T1	T2
N	8	8
Moyenne	3.90	4.13
Médiane	3.90	4.20
Ecart interquartile	0.900	0.900
Etendue	1.80	1.80
Minimum	2.80	3.00
Maximum	4.60	4.80
W de Shapiro-Wilk	0.912	0.926
Valeur p de Shapiro-Wilk	0.365	0.478

Test t pour échantillons appariés

			Statistique	p	Taille de l'effet
T2	T1	W de Wilcoxon	35.0	0.020	Corrélation entre rangs bisériés
					0.944

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.4.1.2 Forte augmentation du SEP(coop) de B sur T1-T2

Statistiques descriptives

	T1	T2
N	10	10
Moyenne	3.90	4.34
Médiane	4.00	4.40
Ecart interquartile	1.05	0.550
Etendue	1.80	1.20
Minimum	2.80	3.80
Maximum	4.60	5.00
W de Shapiro-Wilk	0.915	0.935
Valeur p de Shapiro-Wilk	0.319	0.500

Test t pour échantillons appariés

			Statistique	p	Taille de l'effet
T2	T1	W de Wilcoxon	51.0	0.019	Corrélation entre rangs bisériés
					0.855

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.4.1.3 Différence significative concernant le SEP(coop) de SE sur T1-T4

Statistiques descriptives

	T1	T2	T3	T4
N	30	30	30	30
Moyenne	3.96	4.17	4.22	4.29
Médiane	4.10	4.40	4.60	4.80
Etendue	2.40	3.20	3.00	3.20
Minimum	2.60	1.80	2.00	1.80
Maximum	5.00	5.00	5.00	5.00
W de Shapiro-Wilk	0.944	0.890	0.858	0.810
Valeur p de Shapiro-Wilk	0.118	0.005	< .001	< .001

Friedman

χ^2	ddl	p
15.8	3	0.001

Comparaison pair-à-pair (Durbin-Conover)

Statistiques descriptives

			T1	T2	T3	T4
			Statistique	p		
T1	-	T2	3.097	0.003		
T1	-	T3	2.799	0.006		
T1	-	T4	4.110	< .001		
T2	-	T3	0.298	0.767		
T2	-	T4	1.013	0.314		
T3	-	T4	1.310	0.194		

1.4.1.4 Augmentation modérée du SEP(coop) de SE sur T1-T2

Statistiques descriptives

	T1	T2
N	46	46
Moyenne	4.01	4.15
Médiane	4.20	4.40
Ecart interquartile	1.12	0.950
Minimum	2.60	1.80
Maximum	5.00	5.00
W de Shapiro-Wilk	0.939	0.869
Valeur p de Shapiro-Wilk	0.018	< .001

Test t pour échantillons appariés

			Statistique	p	Taille de l'effet
T2	T1	W de Wilcoxon	520	0.011	Corrélation entre rangs bisériés 0.478

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.4.2 SEC-IAP

1.4.2.1 Très forte augmentation du SEC-IAP de A sur T3-T4

Statistiques descriptives

	T3	T4
N	7	7
Moyenne	4.00	4.86

Statistiques descriptives

	T3	T4
Médiane	4	5
Etendue	2	1
Minimum	3	4
Maximum	5	5
W de Shapiro-Wilk	0.777	0.453
Valeur p de Shapiro-Wilk	0.024	< .001

Test t pour échantillons appariés

			Statistique	p		Taille de l'effet
T4	T3	W de Wilcoxon	15.0	0.048	Corrélation entre rangs bisériés	1.00

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.4.2.2 Forte augmentation du SEC-IAP de SE sur T3-T4

Statistiques descriptives

	T3	T4
N	30	30
Moyenne	3.93	4.43
Médiane	4.00	5.00
Ecart interquartile	2.00	1.00
Etendue	3	3
Minimum	2	2
Maximum	5	5
W de Shapiro-Wilk	0.860	0.714
Valeur p de Shapiro-Wilk	0.001	< .001

Test t pour échantillons appariés

			Statistique	p		Taille de l'effet
T4	T3	W de Wilcoxon	118	0.008	Corrélation entre rangs bisériés	0.735

Note. $H_a \mu_{\text{Mesure 1}} - \mu_{\text{Mesure 2}} \neq 0$

1.4.3 Corrélations entre SEP(coop) et SEC-IAP

1.4.3.1 Forte corrélation positive entre SEP(coop) et SEC-IAP pour A à T4

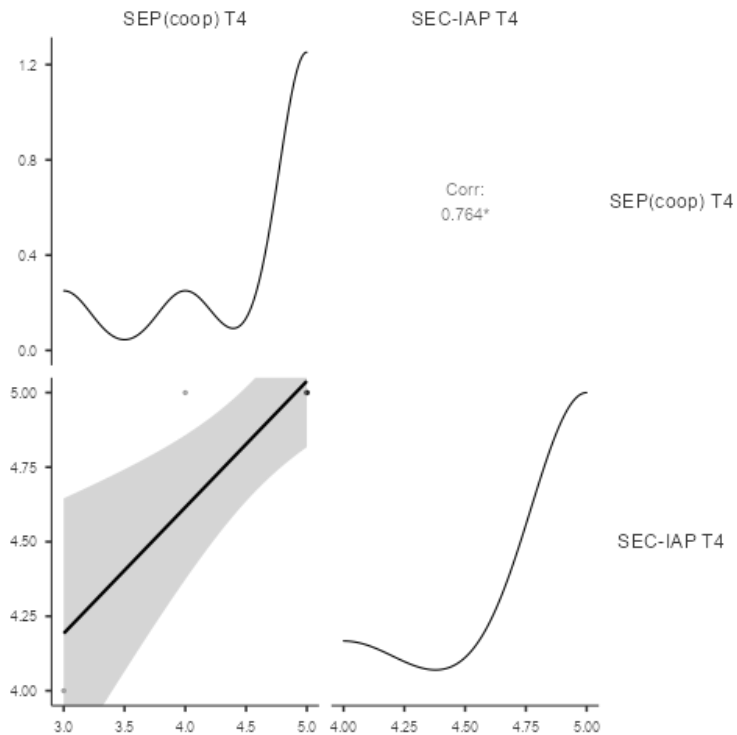
Statistiques descriptives

	SEP(coop) T4	SEC-IAP T4
N	7	7
Moyenne	4.57	4.86
Médiane	5	5
Ecart interquartile	0.500	0.00
Etendue	2	1
Minimum	3	4
Maximum	5	5
W de Shapiro-Wilk	0.646	0.453
Valeur p de Shapiro-Wilk	< .001	< .001

Matrice de corrélation

		SEP(coop) T4	SEC-IAP T4
SEP(coop) T4	Rho de Spearman	—	
	valeur p	—	
	N	—	
SEC-IAP T4	Rho de Spearman	0.764 *	—
	valeur p	0.046	—
	N	7	—

Note. * p < .05, ** p < .01, *** p < .001



1.4.3.2 Forte corrélation positive entre SEP(coop) et SEC-IAP pour C à T4

Statistiques descriptives

	SEP(coop) T4	SEC-IAP T4
N	13	13
Moyenne	4.31	4.31
Médiane	5	4
Ecart interquartile	1.00	1.00
Etendue	3	2
Minimum	2	3
Maximum	5	5
W de Shapiro-Wilk	0.633	0.790
Valeur p de Shapiro-Wilk	< .001	0.005

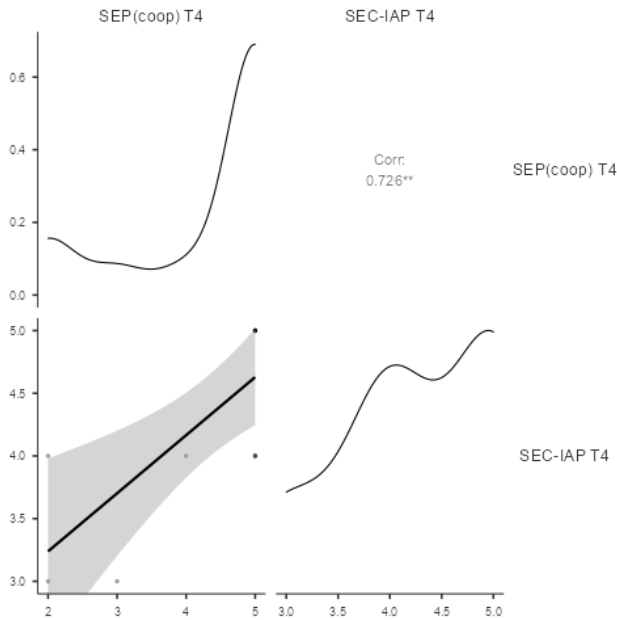
Matrice de corrélation

		SEP(coop) T4	SEC-IAP T4
SEP(coop) T4	Rho de Spearman	—	
	valeur p	—	
	N	—	
SEC-IAP T4	Rho de Spearman	0.726 **	—
	valeur p	0.005	—
	N	13	—

Statistiques descriptives

	SEP(coop) T4	SEC-IAP T4
--	--------------	------------

Note. * $p < .05$, ** $p < .01$, *** $p < .001$



1.4.3.3 Corrélation positive modérée entre SEP(coop) et SEC-IAP pour SE à T4

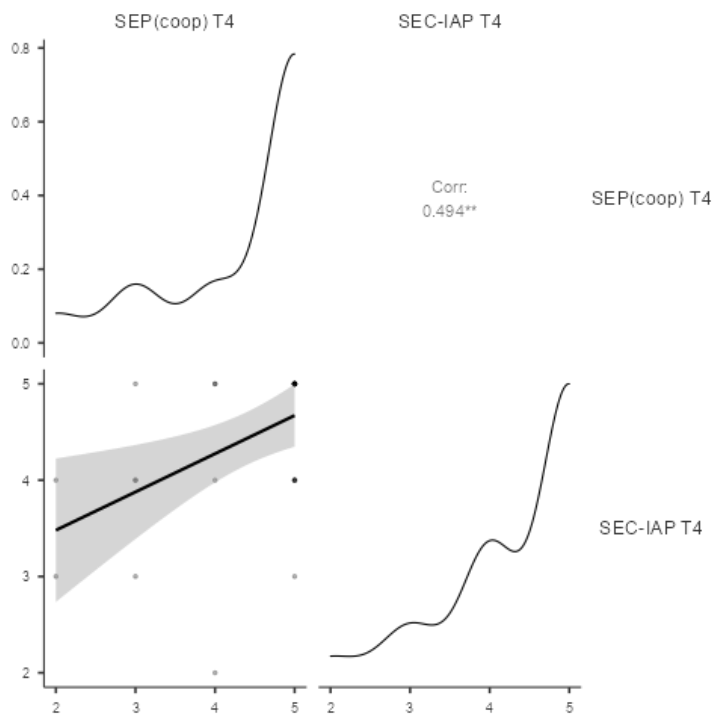
Statistiques descriptives

	SEP(coop) T4	SEC-IAP T4
N	30	30
Moyenne	4.40	4.43
Médiane	5.00	5.00
Ecart-type	0.968	0.817
Etendue	3	3
Minimum	2	2
Maximum	5	5
W de Shapiro-Wilk	0.664	0.714
Valeur p de Shapiro-Wilk	< .001	< .001

Statistiques descriptives

		SEP(coop) T4	SEC-IAP T4
Matrice de corrélation			
		SEP(coop) T4	SEC-IAP T4
SEP(coop) T4	Rho de Spearman	—	—
	ddl	—	—
	valeur p	—	—
	N	—	—
SEC-IAP T4	Rho de Spearman	0.494 **	—
	ddl	28	—
	valeur p	0.005	—
	N	30	—

Note. * p < .05, ** p < .01, *** p < .001



1.4.3.4 Corrélation positive modérée entre SEP(coop) et SEC-IAP pour C à T6

Statistiques descriptives

	SEP(coop) T6	SEC-IAP T6
N	23	23
Moyenne	4.43	4.43

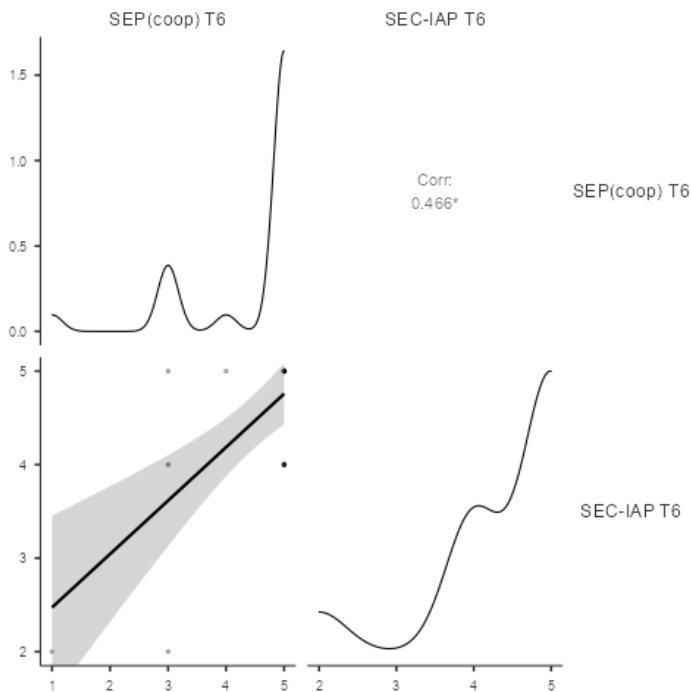
Statistiques descriptives

	SEP(coop) T6	SEC-IAP T6
Médiane	5	5
Ecart interquartile	0.500	1.00
Etendue	4	3
Minimum	1	2
Maximum	5	5
W de Shapiro-Wilk	0.589	0.649
Valeur p de Shapiro-Wilk	< .001	< .001

Matrice de corrélation

		SEP(coop) T6	SEC-IAP T6
SEP(coop) T6	Rho de Spearman	—	
	valeur p	—	
	N	—	
SEC-IAP T6	Rho de Spearman	0.466 *	—
	valeur p	0.025	—
	N	23	—

Note. * p < .05, ** p < .01, *** p < .001



1.5 Hypothèse HO 4

1.5.1 Item 1 : « La coévaluation avec Moodle Atelier m'a paru facile »

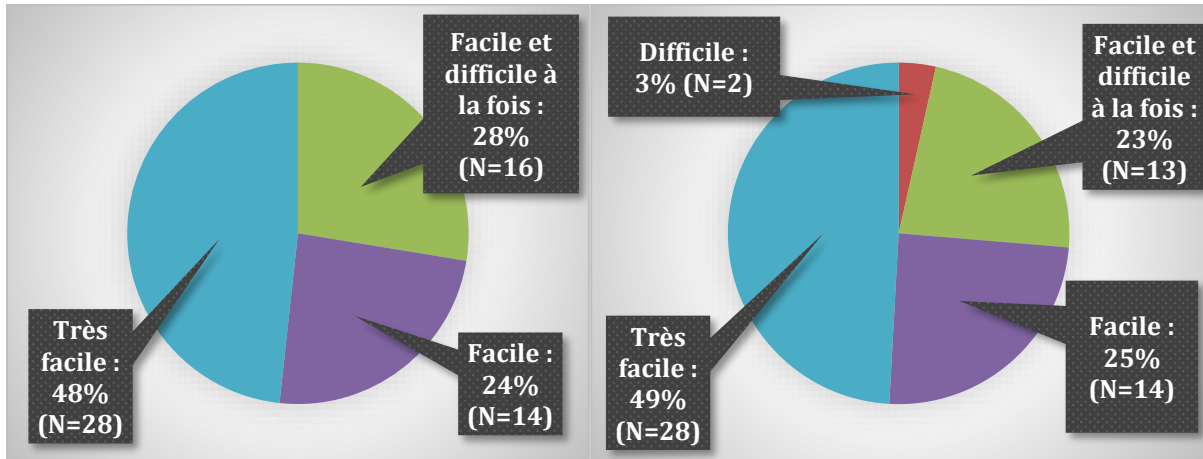


Figure 97 – Item 1 : Analyse descriptive des réponses des élèves à T2 (N = 59) et T4 (N= 57)

1.5.2 Item 2 : analyse catégorielle préliminaire

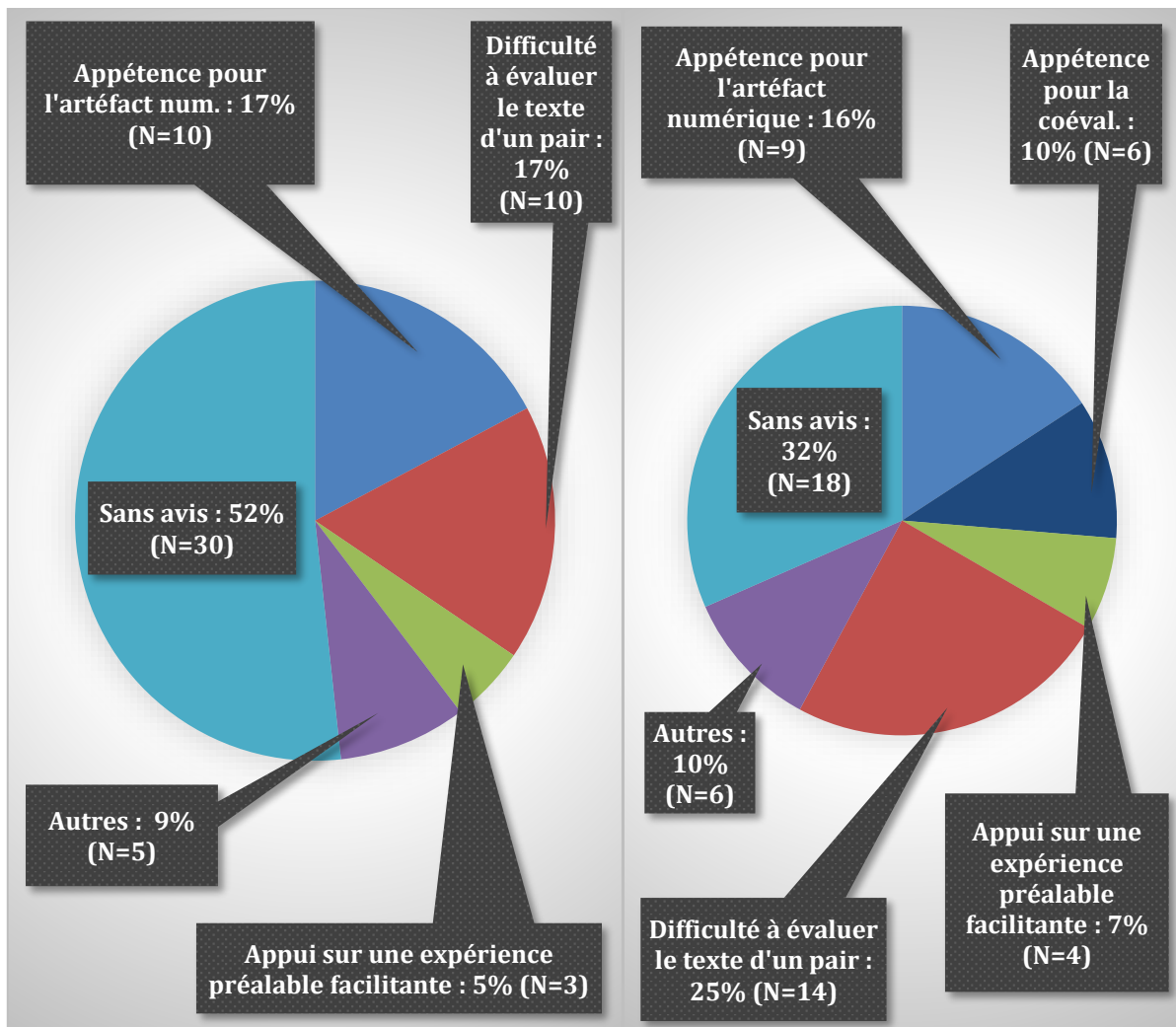


Figure 98 – Item 2 : principales thématiques à T2 (N = 59) et T4 (N = 57)

1.5.3 Item 3 : analyse catégorielle préliminaire

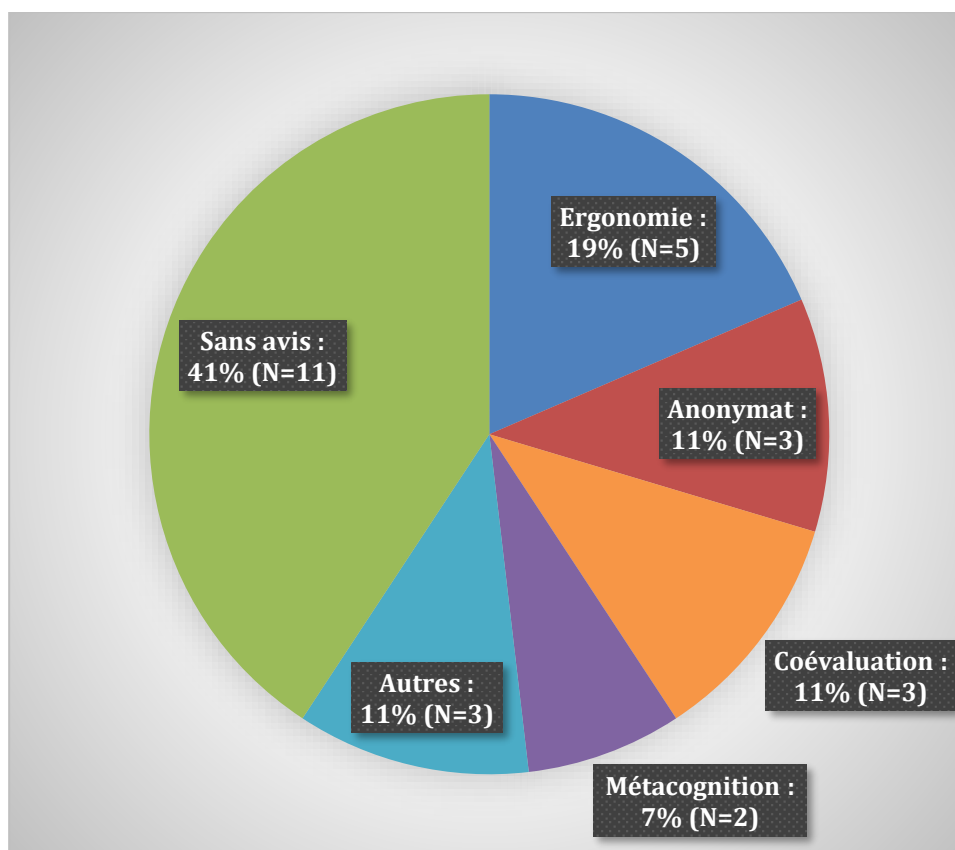


Figure 99 – Item 3 : principales thématiques à T2 (N = 27)

1.5.4 Items 2 et 3 : seconde analyse catégorielle (selon l'approche instrumentale)

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Facile	C'était facile car j'ai vu direct les fautes et c'était agréable de corriger les fautes
Très facile	La coévaluation est plutôt simple, il suffit d'avoir de bonnes bases
Très facile	Facile car dans les ateliers comme cela je suis assez fort et cela me plaît

Tableau 48 – Verbatim d'élèves : présence de prérequis didactiques

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Facile et difficile à la fois	Car je ne suis pas très bon en français
Facile et difficile à la fois	Je n'avais pas compris les questions, j'ai donc eu du mal à y répondre et à corriger le travail de mes camarades
Facile et difficile à la fois	J'avais un peu de mal à noter les inventions des autres élèves car je n'étais pas sûre des relations logiques
Facile et difficile à la fois	C'était pas facile de développer notre texte

Tableau 49 – Verbatim d'élèves : absence de prérequis didactique comme source de conflit instrumental

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Facile et difficile à la fois	Je trouve que c'est compliqué de faire un jugement sur un texte
Facile et difficile à la fois	C'était un peu dur de savoir comment évaluer.
Facile et difficile à la fois	Le travail à faire était facile à comprendre mais au début j'avais un peu de mal à donner mon avis sur le travail de mes camarades
Facile	C'était assez facile, mais assez compliqué à la fois, car on ne sait pas comment bien noter...
Facile et difficile à la fois	Parce que j'ai très rarement évalué d'autres personnes que moi, donc c'était difficile mais d'un autre côté, certaines consignes étaient faciles à évaluer
Facile et difficile à la fois	Je ne savais pas tout le temps quelle case cocher car il y avait des trucs biens et moins biens à la fois.
Très facile	Il n'y a pas de milieu entre « Fragile » et « Maîtrisé », du coup on est obligé de trancher...
Facile et difficile à la fois	Car il faut à la fois être juste et non en même temps
Plutôt facile	L'hésitation entre deux cases
Plutôt facile	C'était bien mais difficile de savoir si notre évaluation est juste
Facile et difficile à la fois	Parce que c'est difficile de savoir mettre les bonnes notes pour évaluer justement un texte

Tableau 50 – Verbatim d'élèves : difficulté à évaluer la production écrite d'un pair (discernement docimologique)

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Très facile	Il y avait beaucoup de fautes d'orthographe ; certaines erreurs étaient catastrophiques...

Tableau 51 – Verbatim d'élève : qualité linguistique insuffisante de certains écrits et difficulté afférente à coévaluer

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Facile	C'était facile car c'était agréable de corriger les fautes
Très facile	J'ai toujours aimé évaluer les autres ; c'est quelque chose que j'adore faire, de base : aider les autres puis les corriger selon les contraintes ; ce n'est pas un souci ou un point où je suis tout à fait faible : j'ai toujours été bon évaluateur plutôt que le faire moi-même le texte, donc je trouve ça mieux que de le créer : j'ai toujours de bons avis personnels qui souvent aident, j'aime aider les personnes qui ont du mal, s'ils en ont en soi vraiment besoin, que ce soit pour un texte ou même une histoire dans la vraie vie
Très facile	Le fait d'évaluer les autres me plait beaucoup
Très facile	C'était mieux et plus intéressant de voir les travaux des autres et de pouvoir les évaluer
Très facile	J'aime surtout pouvoir donner mon avis sur des textes créés par mes camarades
Très facile	Parce qu'il y avait des textes que je comprends
Très facile	J'aime beaucoup voir le niveau des autres et pouvoir donner mon avis sur les textes qu'ils/elles ont écrits
Facile et difficile à la fois	J'ai beaucoup aimé découvrir les inventions des autres élèves

Tableau 52 – Verbatim d'élèves : appétence pour la coévaluation en soi

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Facile et difficile à la fois	Car il faut à la fois être juste et non en même temps
Plutôt facile	Je me sens un peu injuste parce qu'il y a des textes très beaux mais que j'ai assez mal notés (genre 7 ou 9 sur 20)
Facile et difficile à la fois	Parce que c'est difficile de donner une note a une personne, si elle le prend mal
Facile et difficile à la fois	Peur de mal noter
Plutôt facile	Je me sens oppressée

Tableau 53 – Verbatim d'élèves : difficulté à évaluer la production écrite d'un pair (peur de commettre une injustice)

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Plutôt difficile	On n'a pas les mêmes points de vue que les profs
Facile et difficile à la fois	Car je peux aimer le travail et bien l'évaluer et le prof qui va l'évaluer peut moins aimer le travail et mettre une note opposée à la mienne

Tableau 54 – Verbatim d'élèves : difficulté à évaluer la production écrite d'un pair (inter-éval vs. hétéro-éval)

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Facile	Car il est difficile de mettre des commentaires mais sinon je trouve ça plutôt facile
Facile et difficile à la fois	Il faut savoir trouver les bons mots
Facile et difficile à la fois	Car il y a des choses que je n'arrive pas à dire mais que j'ai réussi à répondre
Facile et difficile à la fois	J'ai trouvé difficile le fait de justifier les choix de la coévaluation
Plutôt facile	C'est assez dur de trouver des arguments pour lui expliquer ce qui ne va pas et comment l'améliorer et lui donner une compétence globale pour son travail, pour lui permettre de comprendre et d'améliorer ses textes.
Plutôt facile	Car il faut donner de bonnes explications et trouver la bonne justification
Facile et difficile à la fois	C'est dur de trouver les mots qu'il faut pour évaluer correctement un travail

Tableau 55 – Verbatim d'élèves : difficulté à formuler un écrit commentatif adapté

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Très facile	Car ce n'est pas la première fois que j'évalue une personne
Facile	Ça m'a paru facile car je l'avais déjà fait
Très facile	Car vu qu'on l'avait déjà fait j'ai trouvé ça facile
Très facile	Puisque je l'avais déjà fait la dernière fois cela m'a paru très facile
Très facile	Parce que j'ai l'habitude de corriger mes camarades

Tableau 56 – Verbatim d'élèves : importance de l'expérience acquise

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Facile et difficile à la fois	C'était plutôt long mais ça va c'était simple
Facile	C'est facile à faire
Facile	Simple
Facile	Parce qu'il faut juste évaluer un truc
Très facile	Ce n'était pas très compliqué
Très facile	Parce que c'est facile
Très facile	Personnellement je ne suis pas très indulgent, donc c'était assez facile pour moi
Facile	C'était assez simple, des questions plutôt faciles
Très facile	Parce que c'est un QCM sans mauvaises réponses
Très facile	Oui c'était facile mais je trouve ça très nul : cela va me servir à rien dans ma vie

Tableau 57 – Verbatim d'élèves : contresens concernant l'artéfact pédagogique

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Très facile	J'ai bien aimé l'interface de l'évaluation et l'interaction
Très facile	Car c'est très simple à utiliser
Très facile	Il n'y a pas besoin de faire des textes de beaucoup de lignes : juste cocher
Très facile	C'est facile car ça va vite, c'est facile à corriger et le site est bien organisé
Très facile	Nous avons les critères donc nous nous posons les questions en regardant le texte : il ne reste plus qu'à répondre
Plutôt facile	Moodle est un bon site bien organisé
Très facile	Comment c'est fait me plaît beaucoup
Très facile	J'aime bien l'interface du site, les notations

Tableau 58 – Verbatim d'élèves : facilité d'utilisation

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Très facile	Parce qu'en informatique c'est plus facile
Plutôt facile	En salle informatique
Facile et difficile à la fois	Car on est sur ordi

Tableau 59 – Verbatim d'élèves : appétence pour l'informatique en soi

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Très facile	Parfois le clavier peut amener des erreurs orthographiques, par exemple si on ne sait pas où se trouve l'accent circonflexe

Tableau 60 – Verbatim d'élève : mésusages ou non-usages des interfaces associées à l'artéfact numérique

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Très facile	Ça a été très facile à réaliser car j'ai eu de l'aide du professeur
Très facile	Oui car même si je ne l'avais jamais fait, il y avait des aides
Très facile	Grâce aux exemples vus sur Moodle, j'ai pu évaluer les textes de mes camarades avec sérieux et en sachant ce que je fais.
Facile et difficile à la fois	Car c'est facile vu qu'on a de l'aide
Très facile	On avait des aides

Tableau 61 – Verbatim d'élèves : artéfact numérique et étayage pédagogique

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Facile	Plutôt facile car c'est anonyme donc on a moins peur de donner son avis
Très facile	Elle m'a paru facile car c'était en anonyme

Tableau 62 – Verbatim d'élèves : anonymat

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim élève)
Très facile	J'ai vite compris le travail qu'il fallait faire, sans même avoir eu besoin d'autres explications, et le travail était plutôt bien : moi en tout cas j'ai beaucoup aimé.
Très facile	J'ai vite compris les consignes
Très facile	Grâce aux exemples vus sur Moodle, j'ai pu évaluer les textes de mes camarades avec sérieux et en sachant ce que je fais.
Très facile	Les consignes sont simples et rapides : c'est amusant à faire
Très facile	Les consignes étaient claires et j'ai tout compris donc ça m'a permis de faire tout facilement

Tableau 63 – Verbatim d'élèves : apports de l'artéfact numérique à la compréhension des consignes

1.6 Etayage et performances des élèves : analyse transversale

La coévaluation avec Moodle Atelier m'a paru...	Pourquoi ? (verbatim)	Moyenne scripteur (en %)	Moyenne coévaluateur (en %)
Très facile	Ça a été très facile à réaliser car j'ai eu de l'aide des professeurs.	43	85
Très facile	Oui car même si je ne l'avais jamais fait, il y avait des aides	39	64

Très facile	Grâce aux exemples vus sur Moodle, j'ai pu évaluer les textes de mes camarades avec sérieux et en sachant ce que je fais.	64	76
Facile et difficile à la fois	Car c'est facile vu qu'on a de l'aide	56	81
Très facile	On avait des aides	52	61
Très facile	C'est facile car ça va vite, c'est facile à corriger et le site est bien organisé	64	76
Très facile	Nous avons les critères donc nous nous posons les questions en regardant le texte : il ne reste plus qu'à répondre	28	52
Très facile	J'ai vite compris le travail qu'il fallait faire, sans même avoir eu besoin d'autres explications, et le travail était plutôt bien : moi en tout cas j'ai beaucoup aimé.	62	69
Très facile	J'ai vite compris les consignes	61	47
Plutôt facile	Moodle est un bon site bien organisé	39	64
Très facile	Comment c'est fait me plaît beaucoup	82	63
Très facile	J'aime bien l'interface du site, les notations	77	65
Très facile	Les consignes sont simples et rapides : c'est amusant à faire	53	57
Très facile	Les consignes étaient claires et j'ai tout compris donc ça m'a permis de faire tout facilement	43	85
Plutôt facile	Premièrement c'est grâce aux critères d'évaluation	44	62
	Moyennes	54	67

Tableau 64 – Scores moyens de scripteurs et de coévaluateurs des élèves ayant souligné l'importance de l'étayage

2 Matériel pédagogique des enseignantes partenaires

2.1 T1/T2

2.1.1 Classe A' (4^e)

Écrire une critique

Recherche

1. Comment faire pour écrire une critique ? Quels critères ?
2. Surligner le vocabulaire du théâtre

Mise en commun

- le texte contient des informations : date de création, acteur, décor, thème, intrigue, metteur en scène, dramaturge...
 - décrire ce qu'on voit
 - employer le vocabulaire spécifique au théâtre
- formuler un jugement sur ce qu'on voit (adjectif, verbe d'opinion...)
- forme du texte : paragraphe, texte organisé, connecteurs...
- la critique doit montrer qu'on a bien compris le texte.

2.1.2 Classe C (5^e)

2.1.2.1 Consigne d'écriture

Pour rafraîchir la mémoire, voici le dernier paragraphe du texte :

L'hippogriffe de Roger, sous les ombrages les plus frais et sur l'herbe la plus épaisse, fait cependant tous ses efforts pour fuir, épouvanté de je ne sais quoi d'extraordinaire qui obscurcit tout le bois. Il secoue si violemment le myrte auquel il est attaché qu'il en fait tomber toutes les feuilles sur la terre ; les feuilles en tombant couvrent le pied du myrte tout autour, mais malgré ses plus violents efforts, il ne peut parvenir à s'en détacher.

Consignes pour un écrit projectif :

- L'hippogriffe se débat et tente de s'échapper.
- Dans un texte de 10 à 15 lignes, strictement,

- vous expliquerez ce qui cause la panique de l'animal
- et envisagerez la façon dont il échappe (ou non) au danger,
- sans faire intervenir plus d'un autre personnage en plus de Roger et de l'hippogriffe,
- et sans utiliser de dialogues.

2.1.2.2 Formulaire d'évaluation (APC)

Formulaire d'évaluation ▼

Critères	Niveaux			
Récit projectif, en continuité avec le texte de l'Arioste : au système du présent ; à la 3ème personne.	<input type="radio"/> Insuffisant	<input checked="" type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
Cohérence : même époque, même saison, même lieu. Le caractère des personnages est respecté.	<input type="radio"/> Insuffisant	<input checked="" type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
Le texte est écrit à la façon de l'Arioste : vocabulaire soutenu, moyenâgeux ; tendance à l'exagération.	<input type="radio"/> Insuffisant	<input checked="" type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
Le sens de l'ouïe est mis en valeur.	<input type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input checked="" type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant
Le texte comporte au moins une comparaison.	<input checked="" type="radio"/> Insuffisant	<input type="radio"/> Fragile	<input type="radio"/> Satisfaisant	<input type="radio"/> Très satisfaisant

Figure 100 – Grille de coévaluation pour l'activité de production écrite / étude de la langue menée avec la classe C (5^e) à T1-T2

2.2 T3/T4 : classe C (5^e)



Figure 101 – « Windflowers » de John William Waterhouse (1903) : image support pour la classe C à T3/T4

2.3 T5/T6 : classe C (5^e)

<p>Il me scruta, espérant que je me confie davantage.</p> <p>— Les deux ?</p> <p>— Les deux.</p> <p>— En même temps ?</p> <p>— Ils étaient bien arrivés à créer un enfant ensemble... mourir ensemble, facile !</p> <p>Je fanfaronnais, à mon habitude ; je tentais de dissimuler mes souffrances sous des fables, de la colère, de l'exagération, du sarcasme. Shomintsu mordit à l'hameçon.</p> <p>— Que s'est-il passé ?</p> <p>— Accident de voiture. Mon père conduisait comme une patate. D'ailleurs, mon père faisait tout comme une patate. Moi compris. Il m'a raté. Le seul acte qu'il a réussi, c'est emboutir sa voiture contre un pin parasol en fabriquant des cadavres impeccables, lui et ma mère.</p> <p>— Quand cela s'est-il produit ?</p> <p>— On parle d'autre chose ?</p> <p>Par la suite, chaque fois que Shomintsu revint sur mon enfance, ma famille, mes études, je répondis à l'identique :</p> <p>— On parle d'autre chose ?</p> <p>Ces esquives constantes tarirent le questionnement. Par mon avarice de réponses, je crus avoir convaincu Shomintsu alors que</p> <p>40</p>	<p>— j'allais le découvrir bientôt – je l'avais persuadé que je mentais.</p> <p>Avec les apprentis, j'étudiais du matin au soir, m'initiant aux quatre-vingt-deux prises autorisées et aux coups interdits, luttant avec les plus légers, exécutant de longues courses pour développer la force de mes jambes, travaillant la poussée, l'assouplissement. Mieux que mes camarades, à cause de ma précédente vie de clochard, je supportais la dureté de notre quotidien, l'absence de chauffage au cœur de l'hiver, les vêtements protecteurs refusés aux aspirants, les combats sur la terre battue qui nous rendaient noirs, sales, la sueur collant la poussière à notre peau. Même les humiliations de mes aînés, même l'astreinte à les servir en esclave ne me coûtaient pas. Je cédai au plaisir inédit d'admirer : lorsque Ashoryu, le champion des champions, lequel logeait ailleurs car ses victoires l'avaient enrichi, venait s'entraîner chez nous, je suivais ses gestes avec passion. Si on le comparait à ses adversaires, il n'y avait pas de raisons objectives qu'il l'emporte : il y avait plus volumineux, plus lourd, plus rapide, plus musclé. En revanche, sitôt qu'il montait sur le <i>doyo</i>, il n'y avait pas plus concentré. Il dominait par sa pensée.</p> <p>41</p>
--	---

Figure 102 – Texte support pour la séance T5/T6 avec la classe C (5^e) : *Le Sumo qui ne pouvait pas grossir*, Eric-Emmanuel Schmitt (Albin Michel, 2009)

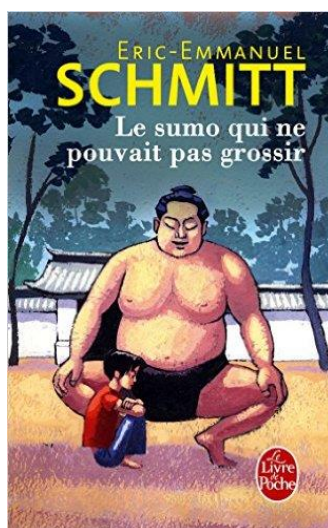


Figure 103 – Aide à l'écriture (séance T5/T6, classe C, 5^e)

2.4 Année N+1

2.4.1 1^{ère} réappropriation

Formulaire d'évaluation ▾

Critère 1
Reprise des informations données dans l'entretien.

Aucune information.
 Il manque de nombreuses informations.
 Il y a la plupart des informations.
 Toutes les informations sont reprises.

Critère 2
Respect du nombre de lignes attendues (10 à 15)

Il y a moins de 10 lignes.
 Le texte comporte de 10 à 15 lignes.

Critère 3
Le vocabulaire est bien adapté et varié.

Le vocabulaire est pauvre et n'est pas adapté.
 Le vocabulaire est varié et bien adapté

Critère 4
Les phrases sont brèves et bien construites.

Les phrases sont mal construites et ne respectent pas la syntaxe
 Il y a quelques maladresses de syntaxe et de ponctuation
 Les phrases respectent la syntaxe

Critère 5
Les conjugaisons sont maîtrisées et les temps bien employés.

Nombreuses erreurs de conjugaisons et d'emplois des temps
 Quelques erreurs
 Les conjugaisons sont maîtrisées et les temps bien employés.

Critère 6
Les accords sont respectés (sujets verbes / adjectifs et noms / participes...)

Nombreuses erreurs
 Quelques erreurs
 Pas d'erreurs

Figure 104 – Formulaire d'évaluation pour la 1^{ère} réappropriation par l'enseignante B (classe B', 3^e, année N+1)

2.4.2 2^{nde} réappropriation

Critère 1 : Le paragraphe comporte dix lignes

- 0 à 4 lignes.
- 4 à 8 lignes.
- 9 à 11 lignes.

Critère 2 : Les informations de la fiche sont toutes présentes.

- Aucune information.
- Quelques informations sont reprises mais il en manque.
- Toutes les informations sont reprises.

Critère 3 : Le vocabulaire est varié et adapté.

- Le vocabulaire est pauvre, répétitif ou d'un niveau de langue inadapté.
- Le vocabulaire est correct mais peu varié.
- Le vocabulaire est varié, riche et très bien adapté au propos.

Critère 4 : Les phrases sont bien construites et respectent la syntaxe.

- De nombreuses erreurs de construction de phrases.
- Quelques maladresses de constructions.
- Des phrases bien construites. La syntaxe est respectée.

Critère 5 : Les accords sujets-verbes sont respectés.

- Plus de 4 erreurs.
- de 1 à 4 erreur(s).
- Aucune erreur.

Critère 6 : Les accords de participes passés et d'adjectifs sont respectés.

- Plus de 5 erreurs.
- Entre 1 et 5 erreur(s).
- Aucune erreur.

Critère 7 : L'orthographe lexicale est respectée.

- Plus de 5 erreurs.
- Entre 1 et 5 erreur(s).
- Aucune erreur.

Figure 105 – Formulaire d'évaluation pour la 2^{nde} réappropriation par l'enseignante B (classe B', 3^e, année N+1)

3 Questionnaire en ligne

Exemple accessible : <https://tripetto.app/run/99EB7XPJDD>

3.1 Information RGPD

Pour rappel, tes données personnelles sont totalement protégées ici — conformément à la réglementation générale européenne sur la protection des données (RGPD). Tes réponses resteront anonymes. Ni ton professeur, ni tes parents, ni tes camarades de classe, ni qui que ce soit d'autre, n'auront accès aux réponses que tu donnes dans le cadre de ce questionnaire de recherche.

3.2 Items d'identification

- Pour commencer, quel est ton prénom stp ?
- Quel est ton nom ?
- Es-tu une fille ou un garçon ?
- Dans quel collège es-tu élève ?
- Dans quelle classe es-tu ?

3.3 Items sur les habitudes évaluatives

- As-tu déjà pratiqué l'auto-évaluation ?
- As-tu déjà pratiqué l'évaluation par les pairs ?

3.4 Items de calibrage

1. Je peux soulever un objet pesant 5 kg.
2. Je peux soulever un objet pesant 10 kg.
3. Je peux soulever un objet pesant 20 kg.
4. Je peux soulever un objet pesant 30 kg.
5. Je peux soulever un objet pesant 40 kg.
6. Je peux soulever un objet pesant 60 kg.
7. Je peux soulever un objet pesant 80 kg.
8. Je peux soulever un objet pesant 120 kg.

3.5 Items SEC

- T3 : La dernière fois, dans quelle mesure t'es-tu senti·e capable d'évaluer avec justesse le travail de tes camarades de classe ?
- T4 : Aujourd'hui, dans quelle mesure t'es-tu senti·e capable d'évaluer avec justesse le travail de tes camarades de classe ?

3.6 Items SEPE

1.	En ce moment, à quel point es-tu capable d'apprendre les mathématiques ?
2.	À quel point es-tu capable d'apprendre la géographie ?
3.	À quel point es-tu capable d'apprendre les sciences ?
4.	À quel point es-tu capable d'apprendre la littérature française ?
5.	À quel point es-tu capable d'apprendre la grammaire française ?
6.	À quel point es-tu capable d'apprendre l'histoire ?
7.	À quel point es-tu capable d'apprendre des langues étrangères ?
8.	À quel point es-tu capable de finir tes devoirs à temps ?

9.	À quel point es-tu capable d'étudier quand il y a d'autres choses intéressantes à faire ?
10.	En classe, à quel point es-tu capable de te concentrer sur le contenu des cours ?
11.	En classe, à quel point es-tu capable de bien prendre les cours en notes ?
12.	À quel point es-tu capable d'utiliser le CDI pour trouver des informations pour les exposés ?
13.	À quel point es-tu capable de t'organiser au jour le jour dans ton travail scolaire ?
14.	À quel point es-tu capable de planifier ton travail scolaire dans le temps ?
15.	À quel point es-tu capable de te souvenir des informations transmises en classe et apprises dans tes manuels scolaires ?
16.	À quel point es-tu capable de t'aménager un endroit où tu puisses étudier sans distractions ?
17.	À quel point es-tu capable de te motiver à faire le travail scolaire ?
18.	À quel point es-tu capable de participer aux échanges en classe ?
19.	À quel point es-tu capable de développer tes compétences sportives ?
20.	À quel point es-tu capable d'apprendre des activités physiques régulières en EPS ?
21.	À quel point es-tu capable d'apprendre les compétences nécessaires pour les sports d'équipe ?
22.	À quel point es-tu capable de résister à la pression, si d'autres collégien·ne·s te poussent à faire des choses qui pourraient t'attirer des ennuis au collège ?
23.	Lorsque tu t'ennuies ou que tu es contrarié·e, à quel point es-tu capable de résister à la tentation de ne pas aller à l'école ?
24.	À quel point es-tu capable de résister à la pression, si d'autres collégien·ne·s te poussent à fumer des cigarettes ?
25.	À quel point es-tu capable de résister à la pression, si d'autres collégien·ne·s te poussent à boire de la bière, du vin ou de l'alcool fort ?
26.	À quel point es-tu capable de tenir tête à quelqu'un qui te demande de faire quelque chose de déraisonnable ou gênant ?
27.	À quel point es-tu capable d'être à la hauteur de ce que tes parents attendent de toi ?
28.	À quel point es-tu capable d'être à la hauteur de ce que tes professeurs attendent de toi ?
29.	À quel point es-tu capable d'être à la hauteur de ce que tes camarades attendent de toi ?
30.	À quel point es-tu capable d'être à la hauteur de ce que tu attends de toi-même ?
31.	À quel point es-tu capable de te lier d'amitié avec des personnes de sexe féminin ?
32.	À quel point es-tu capable de te lier d'amitié avec des personnes de sexe masculin ?
33.	À quel point es-tu capable de tenir une conversation avec d'autres personnes ?
34.	À quel point es-tu capable de travailler en groupe ?
35.	Quand des camarades de classe ne sont pas d'accord avec toi, à quel point es-tu capable d'exprimer tes opinions ?

36.	À quel point es-tu capable de défendre tes intérêts quand tu penses avoir été traité·e de façon injuste ?
37.	Si des personnes t'embêtent ou blessent tes sentiments, à quel point es-tu capable de faire en sorte qu'elles arrêtent ?

3.1 Correspondances items/dimensions SEPE

ITEM #	Libellé (How well can you...)	Sous-échelle de SEP	Traduction
1	learn general mathematics?	SELF-EFFICACY FOR ACADEMIC ACHIEVEMENT	Auto-efficacité pour la réussite scolaire
2	learn geography?	SELF-EFFICACY FOR ACADEMIC ACHIEVEMENT	Auto-efficacité pour la réussite scolaire
3	learn science?	SELF-EFFICACY FOR ACADEMIC ACHIEVEMENT	Auto-efficacité pour la réussite scolaire
4	learn French literature?	SELF-EFFICACY FOR ACADEMIC ACHIEVEMENT	Auto-efficacité pour la réussite scolaire
5	learn French grammar?	SELF-EFFICACY FOR ACADEMIC ACHIEVEMENT	Auto-efficacité pour la réussite scolaire
6	learn history?	SELF-EFFICACY FOR ACADEMIC ACHIEVEMENT	Auto-efficacité pour la réussite scolaire
7	learn foreign languages?	SELF-EFFICACY FOR ACADEMIC ACHIEVEMENT	Auto-efficacité pour la réussite scolaire
8	finish homework assignments by deadlines?	SELF-EFFICACY FOR SELF-REGULATED LEARNING	Auto-efficacité pour l'autorégulation des apprentissages
9	study when there are other interesting things to do?	SELF-EFFICACY FOR SELF-REGULATED LEARNING	Auto-efficacité pour l'autorégulation des apprentissages

10	concentrate on school subjects?	SELF-EFFICACY FOR SELF-REGULATED LEARNING	Auto-efficacité pour l'autorégulation des apprentissages
11	take class notes of class instruction?	SELF-EFFICACY FOR SELF-REGULATED LEARNING	Auto-efficacité pour l'autorégulation des apprentissages
12	use the library to get information for class assignments?	SELF-EFFICACY FOR SELF-REGULATED LEARNING	Auto-efficacité pour l'autorégulation des apprentissages
13	organize your school work?	SELF-EFFICACY FOR SELF-REGULATED LEARNING	Auto-efficacité pour l'autorégulation des apprentissages
14	plan your school work?	SELF-EFFICACY FOR SELF-REGULATED LEARNING	Auto-efficacité pour l'autorégulation des apprentissages
15	remember information presented in class and textbooks?	SELF-EFFICACY FOR SELF-REGULATED LEARNING	Auto-efficacité pour l'autorégulation des apprentissages
16	arrange a place to study without distractions?	SELF-EFFICACY FOR SELF-REGULATED LEARNING	Auto-efficacité pour l'autorégulation des apprentissages
17	motivate yourself to do school work?	SELF-EFFICACY FOR SELF-REGULATED LEARNING	Auto-efficacité pour l'autorégulation des apprentissages
18	participate to class discussions?	SOCIAL SELF-EFFICACY	Auto-efficacité sociale
19	learn sport skills?	SELF-EFFICACY FOR LEISURE TIME SKILLS & EXTRACURRICULAR ACTIVITIES	Auto-efficacité pour les compétences liées aux loisirs et aux activités périscolaires
20	learn regular physical education activities?	SELF-EFFICACY FOR LEISURE TIME	Auto-efficacité pour les compétences

		SKILLS & EXTRACURRICULAR ACTIVITIES	liées aux loisirs et aux activités périscolaires
21	learn the skills needed for team sports (for example, basketball, volleyball, swimming, football, soccer)?	SELF-EFFICACY FOR LEISURE TIME SKILLS & EXTRACURRICULAR ACTIVITIES	Auto-efficacité pour les compétences liées aux loisirs et aux activités périscolaires
22	resist peer pressure to do things in school that can get you into trouble?	SELF-REGULATORY EFFICACY	Auto-efficacité pour l'autorégulation
23	stop yourself from skipping school when you feel bored or upset?	SELF-REGULATORY EFFICACY	Auto-efficacité pour l'autorégulation
24	resist peer pressure to smoke cigarettes?	SELF-REGULATORY EFFICACY	Auto-efficacité pour l'autorégulation
25	resist peer pressure to drink beer, wine or liquor?	SELF-REGULATORY EFFICACY	Auto-efficacité pour l'autorégulation
26	stand firm to someone who is asking to do something unreasonable or inconvenient?	SELF-ASSERTIVE EFFICACY	Efficacité pour s'affirmer
27	live up to what your parents expect of you?	SELF-EFFICACY TO MEET OTHERS' EXPECTATIONS	Auto-efficacité pour répondre aux attentes d'autrui
28	live up to what your teachers expect of you?	SELF-EFFICACY TO MEET OTHERS' EXPECTATIONS	Auto-efficacité pour répondre aux attentes d'autrui
29	live up to what your peers expect of you?	SELF-EFFICACY TO MEET OTHERS' EXPECTATIONS	Auto-efficacité pour répondre aux attentes d'autrui

30	live up to what you expect of yourself?	SELF-EFFICACY TO MEET OTHERS' EXPECTATIONS	Auto-efficacité pour répondre aux attentes d'autrui
31	make and keep female friends ?	SOCIAL SELF-EFFICACY	Auto-efficacité sociale
32	make and keep male friends?	SOCIAL SELF-EFFICACY	Auto-efficacité sociale
33	carry on conversations with others?	SOCIAL SELF-EFFICACY	Auto-efficacité sociale
34	work in a group?	SOCIAL SELF-EFFICACY	Auto-efficacité sociale
35	express your opinions when other classmates disagree with you?	SELF-ASSERTIVE EFFICACY	Efficacité pour s'affirmer
36	stand up for yourself when you feel you are being treated unfairly?	SELF-ASSERTIVE EFFICACY	Efficacité pour s'affirmer
37	deal with situations where others are annoying you or hurting your feelings?	SELF-ASSERTIVE EFFICACY	Efficacité pour s'affirmer

3.1 Codage, dimensions et libellé des items SEPE

RANG	DIMENSION	CODE ITEM	LIBELLE
ITEM 1	Auto-efficacité perçue pour la réussite scolaire	SCOL1	En ce moment, à quel point es-tu capable d'apprendre les mathématiques ?
ITEM 2	Auto-efficacité perçue pour la réussite scolaire	SCOL2	À quel point es-tu capable d'apprendre la géographie ?
ITEM 3	Auto-efficacité perçue pour la réussite scolaire	SCOL3	À quel point es-tu capable d'apprendre les sciences ?

ITEM 4	Auto-efficacité perçue pour la réussite scolaire	SCOL4	À quel point es-tu capable d'apprendre la littérature française ?
ITEM 5	Auto-efficacité perçue pour la réussite scolaire	SCOL5	À quel point es-tu capable d'apprendre la grammaire française ?
ITEM 6	Auto-efficacité perçue pour la réussite scolaire	SCOL6	À quel point es-tu capable d'apprendre l'histoire ?
ITEM 7	Auto-efficacité perçue pour la réussite scolaire	SCOL7	À quel point es-tu capable d'apprendre des langues étrangères ?
ITEM 8	Auto-efficacité perçue pour l'autorégulation des apprentissages	SRL1	À quel point es-tu capable de finir tes devoirs à temps ?
ITEM 9	Auto-efficacité perçue pour l'autorégulation des apprentissages	SRL2	À quel point es-tu capable d'étudier quand il y a d'autres choses intéressantes à faire ?
ITEM 10	Auto-efficacité perçue pour l'autorégulation des apprentissages	SRL3	En classe, à quel point es-tu capable de te concentrer sur le contenu des cours ?
ITEM 11	Auto-efficacité perçue pour l'autorégulation des apprentissages	SRL4	En classe, à quel point es-tu capable de bien prendre les cours en notes ?
ITEM 12	Auto-efficacité perçue pour l'autorégulation des apprentissages	SRL5	À quel point es-tu capable d'utiliser le CDI pour trouver des informations pour les exposés ?
ITEM 13	Auto-efficacité perçue pour l'autorégulation des apprentissages	SRL6	À quel point es-tu capable de t'organiser au jour le jour dans ton travail scolaire ?
ITEM 14	Auto-efficacité perçue pour l'autorégulation des apprentissages	SRL7	À quel point es-tu capable de planifier ton travail scolaire dans le temps ?

ITEM 15	Auto-efficacité perçue pour l'autorégulation des apprentissages	SRL8	À quel point es-tu capable de te souvenir des informations transmises en classe et apprises dans tes manuels scolaires ?
ITEM 16	Auto-efficacité perçue pour l'autorégulation des apprentissages	SRL9	À quel point es-tu capable de t'aménager un endroit où tu puisses étudier sans distractions ?
ITEM 17	Auto-efficacité perçue pour l'autorégulation des apprentissages	SRL10	À quel point es-tu capable de te motiver à faire le travail scolaire ?
ITEM 18	Auto-efficacité perçue pour les apprentissages coopératifs	COOP1	À quel point es-tu capable de participer aux échanges en classe ?
ITEM 19	Auto-efficacité perçue pour les apprentissages coopératifs	COOP2	À quel point es-tu capable de te lier d'amitié avec des personnes de sexe féminin ?
ITEM 20	Auto-efficacité perçue pour les apprentissages coopératifs	COOP3	À quel point es-tu capable de te lier d'amitié avec des personnes de sexe masculin ?
ITEM 21	Auto-efficacité perçue pour les apprentissages coopératifs	COOP4	À quel point es-tu capable de tenir une conversation avec d'autres personnes ?
ITEM 22	Auto-efficacité perçue pour les apprentissages coopératifs	COOP5	À quel point es-tu capable de travailler en groupe ?
ITEM 23	Auto-efficacité perçue pour les compétences liées aux loisirs et aux activités périscolaires	PERI1	À quel point es-tu capable de développer tes compétences sportives ?
ITEM 24	Auto-efficacité perçue pour les compétences liées aux loisirs et aux activités périscolaires	PERI2	À quel point es-tu capable d'apprendre des activités physiques régulières en EPS ?

ITEM 25	Auto-efficacité perçue pour les compétences liées aux loisirs et aux activités périscolaires	PERI3	À quel point es-tu capable d'apprendre les compétences nécessaires pour les sports d'équipe ?
ITEM 26	Auto-efficacité perçue pour l'autorégulation	REG1	À quel point es-tu capable de résister à la pression, si d'autres collégien·ne·s te poussent à faire des choses qui pourraient t'attirer des ennuis au collège ?
ITEM 27	Auto-efficacité perçue pour l'autorégulation	REG2	Lorsque tu t'ennuies ou que tu es contrarié·e, à quel point es-tu capable de résister à la tentation de ne pas aller à l'école ?
ITEM 28	Auto-efficacité perçue pour l'autorégulation	REG3	À quel point es-tu capable de résister à la pression, si d'autres collégien·ne·s te poussent à fumer des cigarettes ?
ITEM 29	Auto-efficacité perçue pour l'autorégulation	REG4	À quel point es-tu capable de résister à la pression, si d'autres collégien·ne·s te poussent à boire de la bière, du vin ou de l'alcool fort ?
ITEM 30	Efficacité perçue pour l'affirmation de soi	FIRM1	À quel point es-tu capable de tenir tête à quelqu'un qui te demande de faire quelque chose de déraisonnable ou gênant ?
ITEM 31	Efficacité perçue pour l'affirmation de soi	FIRM2	Quand des camarades de classe ne sont pas d'accord avec toi, à quel point es-tu capable d'exprimer tes opinions ?

ITEM 32	Efficacité perçue pour l'affirmation de soi	FIRM3	À quel point es-tu capable de défendre tes intérêts quand tu penses avoir été traité·e de façon injuste ?
ITEM 33	Efficacité perçue pour l'affirmation de soi	FIRM4	Si des personnes t'embêtent ou blessent tes sentiments, à quel point es-tu capable de faire en sorte qu'elles arrêtent ?
ITEM 34	Auto-efficacité perçue pour répondre aux attentes d'autrui	ALT1	À quel point es-tu capable d'être à la hauteur de ce que tes parents attendent de toi ?
ITEM 35	Auto-efficacité perçue pour répondre aux attentes d'autrui	ALT2	À quel point es-tu capable d'être à la hauteur de ce que tes professeurs attendent de toi ?
ITEM 36	Auto-efficacité perçue pour répondre aux attentes d'autrui	ALT3	À quel point es-tu capable d'être à la hauteur de ce que tes camarades attendent de toi ?
ITEM 37	Auto-efficacité perçue pour répondre aux attentes d'autrui	ALT4	À quel point es-tu capable d'être à la hauteur de ce que tu attends de toi-même ?

3.2 Sous-échelle SEP(scol)

La sous-échelle SEP(scol) vise à mesurer l'auto-efficacité perçue pour la réussite scolaire, *i.e.* les croyances des élèves en leur capacité à maîtriser des compétences dans différentes disciplines scolaires. Cette sous-échelle se compose de 7 items codés [SCOL] (tableau 70).

CODE ITEM	LIBELLE
SCOL1	En ce moment, à quel point es-tu capable d'apprendre les mathématiques ?
SCOL2	À quel point es-tu capable d'apprendre la géographie ?
SCOL3	À quel point es-tu capable d'apprendre les sciences ?
SCOL4	À quel point es-tu capable d'apprendre la littérature française ?

SCOL5	À quel point es-tu capable d'apprendre la grammaire française ?
SCOL6	À quel point es-tu capable d'apprendre l'histoire ?
SCOL7	À quel point es-tu capable d'apprendre des langues étrangères ?

Tableau 65 – Les 7 items de la sous-échelle SEP(scol)

3.3 Sous-échelle SEP(srl)

La sous-échelle SEP(srl) vise à mesurer l'auto-efficacité perçue pour l'autorégulation des apprentissages, *i.e.* les croyances des élèves en leur capacité à structurer leur environnement pour le rendre propice aux apprentissages, ainsi que leur capacité à planifier et organiser des activités académiques. Cette sous-échelle se compose de 10 items codés [SRL] en référence aux initiales de la terminologie couramment usitée dans la littérature anglosaxonne pour l'autorégulation des apprentissages, *i.e.* « *self-regulated learning* ». Le tableau 71 récapitule ces dix items, leur codage et leur libellé.

CODE ITEM	LIBELLE
SRL1	À quel point es-tu capable de finir tes devoirs à temps ?
SRL2	À quel point es-tu capable d'étudier quand il y a d'autres choses intéressantes à faire ?
SRL3	En classe, à quel point es-tu capable de te concentrer sur le contenu des cours ?
SRL4	En classe, à quel point es-tu capable de bien prendre les cours en notes ?
SRL5	À quel point es-tu capable d'utiliser le CDI pour trouver des informations pour les exposés ?
SRL6	À quel point es-tu capable de t'organiser au jour le jour dans ton travail scolaire ?
SRL7	À quel point es-tu capable de planifier ton travail scolaire dans le temps ?
SRL8	À quel point es-tu capable de te souvenir des informations transmises en classe et apprises dans tes manuels scolaires ?
SRL9	À quel point es-tu capable de t'aménager un endroit où tu puisses étudier sans distractions ?
SRL10	À quel point es-tu capable de te motiver à faire le travail scolaire ?

Tableau 66 – Les 10 items de la sous-échelle SEP(srl)

3.4 Sous-échelle SEP(coop)

La sous-échelle SEP(coop) vise à mesurer l'auto-efficacité perçue pour les apprentissages coopératifs, *i.e.* les croyances des élèves en leur capacité à initier et cultiver des relations sociales et à gérer les conflits interpersonnels. Cette sous-échelle se compose de 5 items codés [COOP]. Le tableau 72 récapitule ces cinq items, leur codage et leur libellé.

CODE ITEM	LIBELLE
COOP1	À quel point es-tu capable de participer aux échanges en classe ?
COOP2	À quel point es-tu capable de te lier d'amitié avec des personnes de sexe féminin ?
COOP3	À quel point es-tu capable de te lier d'amitié avec des personnes de sexe masculin ?
COOP4	À quel point es-tu capable de tenir une conversation avec d'autres personnes ?
COOP5	À quel point es-tu capable de travailler en groupe ?

Tableau 67 – Les 5 items de la sous-échelle SEP(coop)

3.5 Items CI-EIAH

- La coévaluation avec l'atelier Moodle t'a-t-elle paru plutôt facile ou difficile ?
- Pourquoi ?

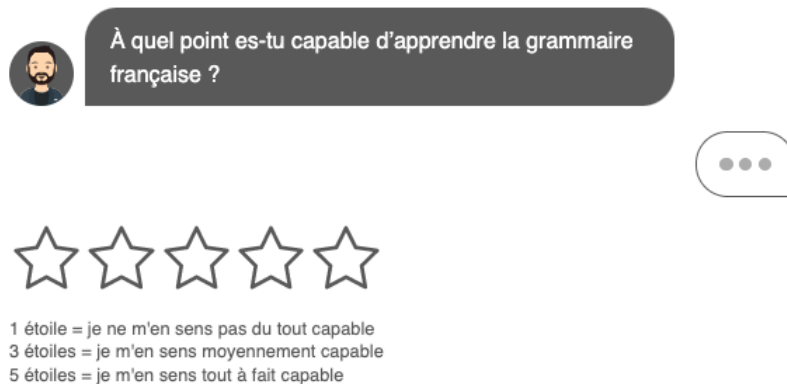
3.6 Items acceptabilité du questionnaire

- Pour toi, ce questionnaire était-il plutôt facile ou difficile à comprendre ?
- Pour toi, l'apparence de ce questionnaire était-elle plutôt agréable ou désagréable ?
- Pour toi, ce questionnaire était-il plutôt pratique ou peu pratique à remplir ?

3.7 Item libre

- Pour finir, as-tu des remarques ou des commentaires à faire ? Ici même les critiques sont les bienvenues ;)

3.8 Captures d'écran



À quel point es-tu capable d'apprendre la grammaire française ?

1 étoile = je ne m'en sens pas du tout capable
3 étoiles = je m'en sens moyennement capable
5 étoiles = je m'en sens tout à fait capable

Detailed description: This is a screenshot of a digital questionnaire. At the top, a dark grey speech bubble contains the question 'À quel point es-tu capable d'apprendre la grammaire française ?' next to a small profile icon of a man with a beard. Below the question, there are five empty star icons arranged horizontally. Underneath the stars, three lines of text provide a legend: '1 étoile = je ne m'en sens pas du tout capable', '3 étoiles = je m'en sens moyennement capable', and '5 étoiles = je m'en sens tout à fait capable'. To the right of the stars, there is a white speech bubble icon containing three dots.

Figure 106 – Echelle de Likert à cinq points avec captions explicatives



Bonjour et merci de participer à ce questionnaire !
Pour rappel, tes données personnelles sont totalement protégées ici — conformément à la réglementation générale européenne sur la protection des données (RGPD). Tes réponses resteront anonymes. Ni ton professeur, ni tes parents, ni tes camarades de classe, ni qui que ce soit d'autre, n'auront accès aux réponses que tu donnes dans le cadre de ce questionnaire de recherche.

Pour commencer, quel est ton nom de famille stp ?

PAUSE

Envoyer

Detailed description: This screenshot shows a questionnaire interface. At the top, a dark grey box contains a welcome message and a detailed RGPD (GDPR) notice. The notice states that data is protected and responses are anonymous, and lists who does not have access to the data. Below this text is a graphic of the European Union flag (a circle of twelve yellow stars on a blue background) with a white padlock icon in the center. Below the graphic, a dark grey speech bubble contains the question 'Pour commencer, quel est ton nom de famille stp ?' next to a small profile icon. To the right of the question is a 'PAUSE' button with a speech bubble icon. Below the question is a long, empty white text input field. At the bottom left, there is a grey button with a paper plane icon and the text 'Envoyer'.

Figure 107 – Information RGPD

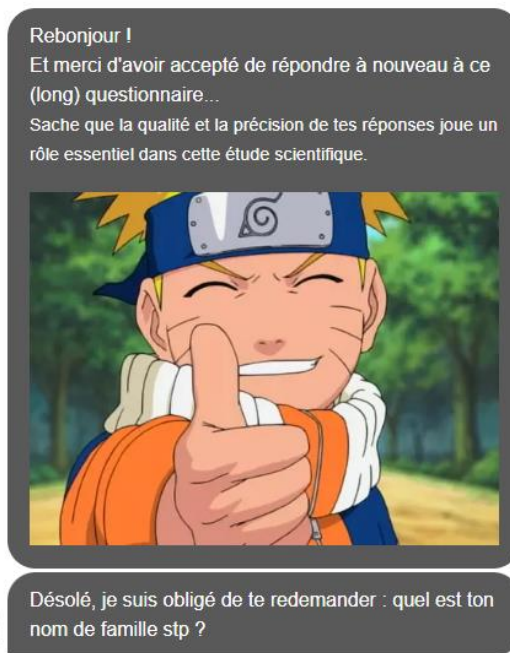


Figure 108 – Exemple de ludification dans Tripetto®

4 Verbatim des retours écrits sur expérience (échantillon C, T6, N = 23)

Elève 1

Cher madame Brulin j'espère que vous allez bien je voulais vous transmettre mon feedback.

ce que j'ai aimé

-j'aime quand l'on me pose des questions avec le questionnaire

-J'aime crée des textes et que d'autres personnes me le notes

-j'aime me placer a la place du professeur en corrigeant les copies de mes camarades surtout en anonyme

-J'aime analyser comment les autre élèves se places a la place des profs car ils n'ont pas autant d'expériences que moi .

ce que je n'est point aimé:

-c'est que madame brulin ne pourras pas avoir nos résultat du questionnaire.

bien cordialement,

Signé : le meilleure élève de France

Réponse de l'enseignante C à l'élève 1

Le meilleur élève, c'est certain.

Elève 2

aligatho madame Brulin

pendant cette évaluation par les pair g aimer

-corriger mes camarade ma bien plus

- travailler sur les pc

-la liberté

ce que je n'ai pas aimer

-le temps pour écrire

-mes notes

-le derouler

merci comme même

Réponse de l'enseignante C à l'élève 2

Vous ne croyez tout de même pas m'échapper ?

PB

Elève 3

Chère Madame Brulin,, je vous transmet mon feedback:

ce que j'ai aimé:

-le fait d'avoir pus corrigé les textes de mes camarades.

-Le fait d'écrire des textes et les partager à mes camarades.

-voir les ressenties de mes camarades envers mon textes

-le questionnaire du début et de la fin.

-Le fait d'avoir connu M. de Khovrine.

Ce que je n'ai pas aimé:

-Les élèves qui noter n'importe comment et sans faire d'efforts.

-et c'est tout.

Cordialement

Réponse de l'enseignante C à l'élève 3

Nous les avons remarqués nous aussi, rassurez-vous !

PB

Elève 4

Chère Mme Brulin

Je voulais vous donner mon avis par rapport à l'atelier d'écriture :

Ce que j'ai aimé :

-Pas grand chose...

Ce que je n'ai pas aimé :

-Je me suis senti injuste car j'ai corrigé des textes beaux mais qui ne respectaient pas les consignes. J'étais horrifié de donner de si mauvaises notes.

-Je trouve que c'est dommage que les correcteurs ne corrigent pas avec les élèves corrigés car les correcteurs peuvent rater des choses que les élèves corrigés avaient pourtant mis.

Recevez, madame, mes sincères salutations distinguées

Réponse de l'enseignante C à l'élève 4

J'ai le même problème quand je corrige, vous avez bien perçu toute la difficulté de l'exercice.

PB

Elève 5

Bonjour,

voici mon avis sur les expériences moodle.

j'ai aimée :

-La liberté d'écriture que nous avons sur notre texte

-L'aide qu'on nous apporte

-La décision que nous avons sur les différents critères choisis en classe

Je n'ai pas aimé :

- La dureté des notes par rapport aux efforts des élèves
- La difficulté a noté objectivement

Bien cordialement

Réponse de l'enseignante C à l'élève 5

Si j'avais dû écrire un feedback, j'aurais écrit exactement le même que le vôtre, ma chère !

PB

Elève 6

Cher Madame Brulin

ce que j'ai aimé :

- avoir une bonne note (15,50) :)
- J'ai bien aimé écrire le texte .
- J'ai bien aimer corriger les textes.

Ce que je n'ai pas aimé :

- Le questionnaire n'est pas animé est trop long
- Et c'est tout bonne journée/bonne soirée

Réponse de l'enseignante C à l'élève 6

Je suis ravie pour votre note, bien méritée !

PB

Elève 7

Chère Madame Brulin,

ce que j'ai aimer :

tout

ce que j'ai pas aimer :

le fragiles alors que mon discours indirect était travaillez

Oh, oh, alors, il faut que nous regardions cela tranquillement ensemble. Il semble me souvenir que c'est une question de disposition et de ponctuation. Le dialogue, c'est le moment où la ponctuation (que vous adorez, n'est-ce pas ;)) a le plus d'importance.

Réponse de l'enseignante C à l'élève 7

Contente que tout vous ait plu par ailleurs.

PB

Elève 8

cher Mme P.Brulin

ce que j'ai aimé :

-Corriger

-Critiquer ext...

Ce que j'ai pas aimé

-avoir une mauvaise note

Réponse de l'enseignante C à l'élève 8

Je vous rappelle que votre texte n'était pas tout à fait fini : on voit ça, et pour Moodle, on laisse comme ça, mais sur votre moyenne, je veux que vos immenses progrès apparaissent !!

PB

Elève 9

bonjour madame Brulin

ce que j'ai aimé :

-les temps d'écriture en cours

-le fait de corriger les copies des autres élèves

-le fait que je puisse voir les commentaires que vous avez fait sur mon texte

ce que je n'ai pas aimé

-le fait que le tableau soit trop strict même si j'ai beaucoup aimé le texte mais que la personne n'a pas respecté les contraintes il aura une mauvaise note

cordialement

passer un bon week-end

Réponse de l'enseignante C à l'élève 9

Nous pouvons encore réfléchir à ce tableau, ou à la note produite par Moodle : trouvons ensemble le meilleur usage.

PB

Elève 10

Chers madame brulin,

j'ai bien aimé dans l'ensemble mais je trouve que la machine note trop sévèrement
Mais la fois que j'ai préféré c quand on a fait le sumo car j'aime bien ce genre de livre
j'ai bien aimé mais mes notes était toujours basse

Réponse de l'enseignante C à l'élève 10

Ravie que le sumo vous plaise. J'en ai offert un exemplaire à Monsieur de Khovrine :
j'espère qu'il aimera aussi.

Dites : c'est la machine qui est trop sévère ou vous qui êtes trop gentil ?? ;) Vous avez
fait d'énormes progrès, vous vous en rendez compte ?

PB

Elève 11

cher Madame Bru lin j'ai beaucoup apprécié cette séance surtout >>>noter les gens
pour les améliorer pour le prochain texte >et aussi le fait de voir le travail des gens

Réponse de l'enseignante C à l'élève 11

Je suis contente que vous perceviez le but ultime, qui est de s'améliorer. Nous finirons
votre texte à vous.

PB

Elève 12

cher Madame Brulin je trouve que le projeté moodle et une super expérience se-la ma
permis de pouvoir apprendre a corriger j'ai adoré tout faire pour moi il n'y a aucun point négatif
j'aime réé beaucoup recommencer en 4e

cordialement

Réponse de l'enseignante C à l'élève 12

Nous proposerons cet outil à votre futur professeur l'an prochain alors ;) Il faudra
m'aider !

PB

Elève 13

Chère Mme Brulin, ce qui m'a plu avec moodle c'est:

-l'anonymat

-pouvoir corriger (se mettre à la place des professeurs),

ce qui m'a pas plu c'est:

-devoirs choisir entre satisfaisant et fragile

Merci bonne après midi

Réponse de l'enseignante C à l'élève 13

C'est un choix difficile : peut-être manque-t-il une rubrique ? Peut-être qu'on aurait le même problème avec plus de rubriques ?

;))

PB

Elève 14

Chère madame Brulin

Ce que j'ai bien aimée sur moodle est:

-De corriger les copies d'autres personnes.

-Et d'être avec monsieur de Khovrine

Réponse de l'enseignante C à l'élève 14

Je vous rejoins sur le fait que Monsieur de Khovrine est de bien agréable compagnie !

PB

Elève 15

bonjour Mme Brulin j'ai bien aimer les ateliers d'écriture. J'ai également bien aimais pouvoir évaluer les autre texte.

Réponse de l'enseignante C à l'élève 15

Les ateliers d'écriture, c'est ce que je préfère. Vous vous souvenez de ceux de l'an dernier ?

PB

Elève 16

cher Mme Brulin j'ai bien aimé les trois séance de moodle car j'aime bien quand on évalue nos camarades et aussi quand on laisse des feedback cordialement

Réponse de l'enseignante C à l'élève 16

C'est très instructif de voir ce que les autres ont écrit. Je me réjouis toujours de vous lire tous (moins de vous noter, j'avoue !)

PB

Elève 17

je trouve que l'eval et bien car on peut voir les feed back et ce que les autre pense de notre texte

Réponse de l'enseignante C à l'élève 17

C'est ainsi qu'on progresse, en ayant les avis des autres. Des avis croisés, pas d'une seule personne !

PB

Elève 18

Aligato madame sans plus tarder je vous cite la liste de mes préférence et de mes remarques

ce que j'ai aimé :

- corriger les textes
- travailler sur les pc

ce que j'ai pas aimé :

- ma note
- le concept
- le temps perdu a mon gout sur ce projet
- le manque de temps pour écrire notre texte d'où ma note
- les questionnaire trop barbant et répétitif

désolé pour l'honnêteté cash

Réponse de l'enseignante C à l'élève 18

C'est une de vos très grandes qualités, l'honnêteté. Il n'y a rien à pardonner, au contraire, merci de cette franchise. Nous en parlerons avec monsieur de Khovrine.

Au fait, la dernière heure consacrée à l'écriture, ne l'avez-vous pas passée à discuter avec un camarade ?

PB

Elève 19

Cher madame Brulin la seule et unique prof de français, Sa majesté Nicolas nous à demander vous vous écrire un texte pour savoir se qui nous a plu et déplu (notre avis):

Ce qui m'a plus :

- corrige les copies de nos camarade en particulier
- plusieurs sujet différent
- divers texte avec chacun son imagination.
- ne pas savoir qui on corrige pour ne pas faire de préférence

Ce qui m'a déplu:

- l'anonymat(juste pour ma copie)
- les questionnaire un peu trop long
- pas assez de temps pour corrigé

Cordialement,

Votre élève préférée [avec l'élève 21] :)

Réponse de l'enseignante C à l'élève 19

La question du temps est un problème dont nous discutons avec Monsieur de Khovrine, en effet.

Comment diable avez-vous deviné que vous étiez mes préférées ;))

PB

Elève 20

Cher Mme Brulin, voici mon avis sur les 3 ateliers que nous avons pu faire en compagnie de Mr Nicolas.

J'ai bien aimé l'interface de Moodle qui nous as permis de corriger nos camarades.

Mais en revanche je n'ai pas trop aimé mettre une note car celle là n'est pas forcément juste pour le travail qui a été fournis par la personne en question. En revanche j'ai bien aimé voir les commentaires qui ont été fais par mes camarade et voir comment ils ont corrigés mon travail.

Bonne journée,

Cordialement

Réponse de l'enseignante C à l'élève 20

Corriger, c'est une tâche difficile : et moi qui espérait vous laisser dorénavant cette corvée !!

J'aime bien, comme vous, l'interface !

PB

Elève 21

Chère madame Brulin, voici mon avis sur les ateliers que nous avons pu faire au cours de l'année avec sa majesté Nicolas. J'ai beaucoup aimé le fait de pouvoir découvrir comment nos camarades ont imaginés leurs textes/suites de textes, car nous n'avons pas tous les mêmes idées et c'est génial de pouvoir les partager. En revanche je trouve ça un peu compliqué car parfois il y a beaucoup de critères à traiter et pas assez de temps donc nos évaluations peuvent parfois être « bâclées ».

Globalement j'aime beaucoup cette méthode et je serai ravie de recommencer ;)

Réponse de l'enseignante C à l'élève 21

Chère élève préférée avec [l'élève 19] !

Oui, vous n'imaginez pas le temps que je peux passer sur une copie : c'est même un vrai problème. Je suggérerai à votre enseignant de l'an prochain que vous puissiez vous resservir de Moodle !

PB

Elève 22

Cher Mme Brulin voici mon avis sur les 3 séances moodle que nous avons pu faire en compagnie de Mr Nicolas .

-J'ai bien aimé évaluer les autres camarades de classe

-j'ai pas aimer les notes que les camarades mon mit

Réponse de l'enseignante C à l'élève 22

Et pourtant, vous avez vraiment bien évolué. On en reparle : je vous ai mis une appréciation très très positive ce trimestre !!

PB

Elève 23

Chère Mme Brulin,

Ce qui m'a plu pendant cet atelier moodle c'est :

- corriger des copies de nos camarades sans qu'ils sachent que c'est nous.
- pouvoir mettre un commentaire

Bonne journée !!

Réponse de l'enseignante C à l'élève 23

Pour certaines choses, il vaut bien mieux garder l'anonymat ! ;))

PB

5 Transcripts des focus groupes (C, T6)

5.1 Transcript du 1^{er} focus groupe (C1, T6, N = 12)

01:50 Elève 1

Je voulais dire que le temps d'écriture n'était pas assez long : on n'avait pas le temps de faire un texte complet et bien approfondi et tout ce qu'on voulait faire dedans. Voilà.

05:08 Elève 2

Je trouvais que comme par exemple parce que on a tous été correcteur j'ai lu des très très beaux textes mais qui ne respectaient pas les consignes, du coup j'étais vraiment désolé, je me sentais un peu coupable de mettre des si mauvaises notes à de très très beaux textes. Et aussi par exemple quand on corrige un texte on n'est pas attentif à tout ce qui est dans le texte : on rate forcément certaines choses et donc du coup ça serait peut-être bien que celui qui se fait corriger nous dise si on trouve pas par exemple des comparaisons, qu'ils nous disent que par exemple là il y en a une et que du coup on puisse un peu améliorer sa note.

05:54 Chercheur-intervenant

Très intéressant ça. Il faudrait une espèce de droit de réponse ?

05:57 Elève 2

Oui.

05:59 Chercheur-intervenant

Pouvoir dire : « Ah, là je ne suis pas d'accord ». Très intéressant, ça.

06:02 Elève 2

Dans l'anonymat, parce que si on se met à se crier dessus, on n'aura plus envie de corriger. Donc un droit de réponse mais dans le secret, pour que ça ne soit pas l'anarchie.

07:05 Elève 3

Aussi moi je trouve que ça serait bien plutôt de rajouter une note genre à la fin, pas forcément bonus mais juste une case où on puisse rajouter du feedback qui peut donner du bonus, parce qu'il y a parfois des beaux textes qui sont très bien faits mais ne respectent pas forcément les consignes, et du coup ils perdent des points alors que le texte, en soi, on s'est donné du mal à le faire.

07:46 Chercheur-intervenant

Alors je me permets de vous relancer sur cette idée de feedback car ça m'intéresse beaucoup : donc vous voudriez un moyen pour valoriser certaines copies qui ne sont pas forcément dans les critères du tableau mais qui méritent par leur effort. Et quel serait ce moyen ?

08:16 Elève 4

Peut-être un feedback hors sujet ? Pas le feedback que tu corriges...

08:28 Elève 5

Quelque chose qui n'est pas obligatoire...

08:34 Elève 3

Ah quand on réévalue il y a un truc qui apparaît un feedback bonus ou genre par exemple c'est sur trois par rapport à une note bonus, ce qu'on pense nous du texte, enfin... En gros, qu'il y ait pas des consignes spécialement dans le feedback bonus, mais que nous on puisse rajouter quelques points si le texte on voit qu'il a eu du mal à le faire et qu'au final il a pas trop respecté les consignes, mais il a quand même des points.

13:06 Elève 6

Ce que j'ai aimé c'est le fait de pouvoir corriger des textes que d'autres personnes ont faits, le fait de partager nos avis, nos sentiments sur les choses que les autres ont faites, sur ce qu'on a aimé ou pas.

13:55 Elève 7

Moi j'ai bien aimé les thèmes sur lesquels on devait écrire.

14:12 Elève 8

Moi, je trouve que c'était bien de faire le correcteur.

14:22 Chercheur-intervenant

Pourquoi ?

14:23 Elève 8

Parce que c'est un peu se mettre à la place de la prof et savoir à quoi elle est confrontée. Et aussi de voir tous les styles de textes des gens, même si c'est anonyme.

16:08 Elève 9

Juste, je voulais savoir pourquoi les correcteurs sont anonymes. Parce que je pense pas quand même qu'on va taper si on a une mauvaise note. Ça serait bien en fait de savoir qui nous a corrigé. Je pense qu'on pourrait lever l'anonymat après, pour pouvoir se concerter avec lui et savoir pourquoi il a mis cette note.

17:42 Elève 10

Je suis pas sure, mais déjà c'est mieux parce que comme ça au lieu des messages anonymes on peut aller parler à la personne, enfin... sans lui taper dessus ; mais on peut aller parler à la personne en face, mais calmement et juste lui demander pourquoi il a dit « insuffisant », pourquoi il a dit « satisfaisant ». Comme ça si on peut refaire après un retour sur le texte ça permettrait de changer après avoir discuté avec la personne et de voir si les arguments sont valables ou pas. Ça serait bien de faire, regarder la note, etc., discuter avec les autres et revoir notre texte.

18:24 Elève 9

Bah tu le menaces : tu lui dis je te mets un coup de couteau il va te mettre une meilleure note...

18:28 Elève 8

Non mais sans pouvoir changer la note après...

18:32 Elève 10

Si les arguments sont valables...

18:35 Elève 11

Ça sera jamais valable...

18:36 Elève 10

Si j'ai pas vu un verbe que t'as mis, eh ben tu vas me dire : « Oui, mais il était là... ». Je vais dire « Ah, oui... ». Du coup je vais faire un retour et je vais te rajouter un point.

18:42 Elève 12

Oui, si quelqu'un peut le prouver... Mais si tu dis un argument en mode « Elève 9 »...

18:50 Elève 10

C'est pas un argument... Non mais j'ai donné un argument, genre : élève 6 elle a noté élève 8, c'est faux : personne ici pourra prouver que l'argument il est faux si t'améliores ma note parce qu'on est amis, on améliore si l'argument il est valable.

19:46 Elève 7

Faudrait cocher une option pour qu'on vote, enfin pour que l'autre il sache qui l'a évalué.

5.2 Transcript du 2nd focus groupe (C2, T6, N = 11)

00:07 Elève 1

Je n'ai pas aimé ma note d'évaluatrice, mais c'est de ma faute si j'ai eu cette note-là.

00:37 Elève 2

Moi non plus je n'ai pas aimé ma note.

01:00 Elève 3

Des fois il y a beaucoup de critères qu'il faut chercher dans les textes des autres etc., du coup c'est un peu compliqué et en une heure on n'a pas forcément toujours le temps, du coup, de bien trouver chaque critère, et du coup souvent on loupe des choses, enfin on rate des choses ; on dit, on met « insuffisant », enfin donc etc., alors qu'on n'a pas assez cherché.

02:00 Elève 4

Surtout je voulais reprendre sur ce qu'a dit l'élève 3 c'est surtout le truc qu'en fait en 1h c'est vrai qu'on n'a pas le temps parce qu'on a par exemple le texte de l'élève 5 c'est pas moi qui l'ai corrigé mais quand je l'ai vu tout à l'heure à son ordinateur j'ai vu qu'il était très long donc je pense que la personne qui a dû corriger elle a dû mettre du temps pour lire et tout

02:30 Elève 5

C'est pour ça que j'ai eu 7 ! (rire général)

02:55 Elève 6

Si un moment les notes vont compter dans la moyenne et que c'est les élèves qui ont corrigé, et que les élèves n'ont pas forcément eu une bonne note de coévaluateur, du coup ça peut rabaisser notre moyenne, alors que c'est pas forcément juste.

03:20 Chercheur-intervenant

Je reformule : vous me dites si c'est ça. « Sentiment d'injustice si certains élèves évaluateurs se sont trompés et que la note compte dans la moyenne ».

Alors là je réponds juste à ce que vient de dire l'élève 6 parce qu'il y en a peut-être d'autres qui se posent la question.

Effectivement on a eu le cas tout à l'heure. L'élève 1 me montrait une évaluation où clairement on voyait qu'il y avait trois élèves évaluateurs qui donnaient – je ne sais plus pour quel critère c'était : pour le conditionnel ou le dialogue au discours direct – « très satisfaisant » ou « satisfaisant » et puis il y avait une évaluation qui était complètement décalée et qui disait que c'était « insuffisant ». Donc comment on peut passer de « très satisfaisant » à « insuffisant » ? Et donc on en a discuté avec l'élève 1 et on a dit : le problème c'est que l'élève qui a évalué ici, peut-être que lui-même ne savait pas que c'est qu'un dialogue maîtrisé. Et du coup, il a mis « insuffisant », mais il s'est un peu trompé. Donc, lui ou elle n'a pas eu une bonne note de coévaluateur sur ce coup-là. Mais vous, en fait, ça ne vous pénalise pas vraiment. Pourquoi ?

05:40 Elève 3

La prof nous avait dit que les notes que les autres nous mettaient ne comptaient pas parce qu'ils pouvaient se tromper et que c'était pas juste.

05:50 Chercheur-intervenant

Alors en fait les notes comptent mais la note de votre enseignante a un coefficient tellement énorme que c'est elle qui va décider de la note. L'influence des autres notes données est minime. L'essentiel pour vous ici c'était le fait d'évaluer les copies des copains : normalement c'est censé vous aider.

07:20 Elève 2

Un truc que je trouve dommage c'est qu'on peut pas savoir qui c'est qu'on évalue. Il y a des textes que je trouve très bien mais je peux pas savoir qui c'est que c'était. Et pareil pour des gens qui trouvaient que mes textes étaient par exemple très nuls, je peux pas savoir qui c'est. Des fois dans le feedback c'est tout à fait clair : on comprend pourquoi il nous a mis ça

et pourquoi il a compris ça ; comme des fois on met une mauvaise note et dans le feedback on met « c'est nul ».

08:37 Chercheur-intervenant

Donc le feedback insuffisant parfois pour justifier, donc un dialogue avec l'évaluateur serait nécessaire pour expliquer.

08:56 Elève 7

Moi je trouve que c'est mieux de rester dans l'anonymat, parce qu'après peut-être que si on sait les prénoms on va être énervé ou je sais pas quoi, ça crée des conflits.

10:25 Elève 3

Moi, ce que j'ai bien aimé, c'est que des fois, par exemple quand on donne une consigne pour écrire un texte, on sait pas comment les autres imaginent. On n'imagine pas pareil que les autres, comment on va faire le texte etc. Par exemple, moi, quand j'ai écrit mon texte sur « Le Sumo qui ne pouvait pas grossir », je me suis mis dans la peau d'Ashoryu alors qu'il y en a certains qui se sont mis dans la peau de Jun. Et du coup, c'est bien quand on peut lire les textes des autres, parce qu'on peut savoir comment eux ils ont imaginé la scène.

10:50 Chercheur-intervenant

Pour l'inspiration, en quelque sorte ?

10:52 Elève 3

Oui, voilà.

11:10 Elève 4

Moi je trouve que ça a été bien de garder l'anonymat parce que ça a évité de faire du favoritisme. Par exemple, comme l'élève 7 c'est mon amie, si j'avais eu son texte, bien sûr moi je lui aurais mis une bonne note. Comparé à ce que ce soit par exemple l'élève 2 (rire général).

11:58 Elève 1

Je suis d'accord avec l'élève 4 : si on avait vu les prénoms des textes, je trouve qu'on aurait fait du favoritisme ; mais ça aurait été bien par exemple de voir, mais que à la fin.

12:16 Elève 2

Mais voilà c'est ça que je disais !

6 Entretiens semidirigés enseignantes (liens pérennes renvoyant vers les enregistrements audios)

6.1 Enseignante A (T2)

Lien complet :

<https://api.nakala.fr/embed/10.34847/nkl.b8bc9032/53ca6d42fe351e4235dbdae2df3cc58b845bbcad>

Lien court (non pérenne) :

<https://page.hn/ql8r7h>

6.2 Enseignante B (T4)

Lien complet :

<https://api.nakala.fr/embed/10.34847/nkl.b8bc9032/9dee14e77747e61877093723e19f13ae2ec7123e>

Lien court (non pérenne) :

<https://page.hn/4s0zwt>

6.3 Professeure-documentaliste du collège C (T4)

Lien complet :

<https://api.nakala.fr/embed/10.34847/nkl.b8bc9032/f2d2573ac82552714c3f05a4283e47a6d723f7a6>

Lien court (non pérenne) :

<https://page.hn/0u2z9b>

6.4 Enseignante C (T6)

Lien complet :

<https://api.nakala.fr/embed/10.34847/nkl.b8bc9032/64a5ed81be97f04b4004ddbe9c937899b6f75111>

Lien court (non pérenne) :

<https://page.hn/z992tw>

6.5 Enseignante B (année N+1)

Lien complet :

<https://api.nakala.fr/embed/10.34847/nkl.b8bc9032/62eb68771d173098f327536f733e69786b2c472a>

Lien court (non pérenne) :

<https://page.hn/wwvq65>

7 Focus groupe enseignantes A, B et C (T7) : lien pérenne renvoyant vers l'enregistrement audio

Lien complet :

<https://api.nakala.fr/embed/10.34847/nkl.b8bc9032/7065ccb98950533f55390f1a80c7028308c1955b>

Lien court (non pérenne) :

<https://page.hn/hgj7cw>

8 Checklist Moodle Atelier



Catachrèse « Coévaluation »

8.1 Préparation par l'enseignante et les élèves

- ☒ **Consigne** d'écriture & **critères** de réussite (dispositif semble aider à une reformulation + explicite)
- ☒ **Réflexion docimologique** avec élèves : oui mais bref et générique car manque de temps (pragmatisme)
- ☒ **Descriptifs** co-construits pour chaque critère de réussite (voir ci-dessus)

8.2 Configuration de l'atelier en amont

- ☒ Informations générales
- ☒ Réglages de remise des travaux
- ☒ Modalités d'évaluation (auto-évaluation : non)
- ☒ Feedback entre élèves : requis
- ☒ Travaux exemplaires : non (pour ne pas biaiser les résultats, mais oui avec 5^e4)
- ☒ Retirer localement la permission « Voir les auteurs » pour les « Etudiants »
- ☒ Réglages d'évaluation : par critères, note sur 20
- ☒ Préparer le formulaire d'évaluation
- ☒ Achever la phase de mise en place et passer à la phase de remise
- ☒ Ajouter « Ressources » liens URL vers sondages de début et de fin
- ☒ Attribuer localement le rôle « Etudiant » à l'enseignante
- ☒ Déterminer le nombre d'évaluations par élève : 4, auto-évaluation non incluse
- ☒ Attribuer les travaux à l'enseignante (après les avoir attribués aux élèves)
- ☒ Pondération : affecter coeff. 16 pour évaluations par l'enseignante

8.3 Manipulations sur l'atelier en aval

- Calculer les notes
- Recalibrer si nécessaire le degré d'exigence de l'algorithme : + indulgent ? + strict ?

9 Lettre type d'information aux parents



Prénom & nom de l'enseignant
Coordonnées du collège d'exercice



Année scolaire 2021-2022

Madame, Monsieur, chers parents d'élèves,

Dans le cadre du cours de français, les élèves des classes de ... et ... vont prendre part à une **expérimentation pédagogique**, fruit d'un partenariat avec un professeur de français et chercheur, Nicolas de Khovrine, qui prépare une thèse de doctorat en sciences de l'éducation, en lien avec notre chef d'établissement M. ..., les inspecteurs de Lettres et les CARDIE (Cellule Académique de Recherche et Développement pour l'Innovation et l'Expérimentation) des rectorats de Paris et Montpellier.

Cette étude s'intéresse au **numérique éducatif** en tant qu'outil au service des collégiens et de leurs enseignants, pour l'amélioration des apprentissages en français. Concrètement, grâce à un logiciel libre et gratuit accessible directement via l'ENT du collège, les élèves vont avoir la possibilité de **s'évaluer les uns les autres** de manière anonyme, constructive et bienveillante, sous ma supervision. Cette approche vise à **renforcer les apprentissages**, mais aussi la **confiance en soi** et l'**esprit de coopération**.

Ce travail sera mené en **conformité avec les Programmes officiels** de cycle 4, dans le **respect de la progression annuelle** que j'ai prévue.

Surtout, les plateformes et logiciels utilisés seront l'ENT et Moodle, c'est-à-dire des **espaces numériques à accès restreint et sécurisé**. Conformément au Règlement général pour la protection des données (**RGPD**), les données collectées dans le cadre de cette expérimentation seront intégralement anonymisées.

Merci de bien vouloir compléter le coupon ci-dessous, attestant que vous avez bien pris connaissance de cette expérimentation pédagogique et y consentez de manière libre et éclairée.

Bien cordialement,

(Nom de l'enseignant)

Je soussigné(e), (NOM & Prénom), responsable de l'élève : (NOM & Prénom), classe de atteste par la présente avoir été informé(e) de l'expérimentation pédagogique menée en classe de français autour du numérique éducatif, à laquelle je consens de manière libre et éclairée.

Date et signature :

10 Article coécrit avec les 3 enseignantes partenaires

Une recherche collaborative sur l'évaluation coopérative

En 2021-2022, trois enseignantes, un chercheur et 63 collégiens de cycle 4 ont, au moyen du logiciel Moodle Atelier, évalué conjointement des productions écrites ciblant l'étude de la langue. La notion de groupe était au cœur de cette « coexpérimentation pédagonumérique ».

Favoriser le travail en groupe des élèves

La première étape est l'élaboration collective des critères de réussite que chaque élève va appliquer à son travail et à celui des autres : il les conscientise et se les approprie. Ensuite, seul devant son ordinateur, chacun fait cependant groupe avec la classe. Au départ, l'accès à l'activité est parfois fastidieux, mais cela incite à l'entraide et favorise ainsi un apprentissage entre pairs de la littératie numérique.

Le collectif entre en jeu par le regard porté sur le travail des autres, dont on se nourrit avant de revenir vers le sien. La rédaction des feedbacks invite les élèves à justifier leurs remarques ou à valoriser les réussites dans une démarche bienveillante. Les correcteurs se montrent souvent plus indulgents qu'on ne le pourrait croire, hésitent à sanctionner oublis et maladresses. Des nuances docimologiques⁷² s'affinent progressivement, clarifiées par le collectif en amont de la coévaluation.

Cette rédaction amène un travail métacognitif : les remarques adressées aux autres sont des conseils que l'on se donne à soi-même. Cette démarche est particulièrement efficace pour la construction des compétences : il ne s'agit pas seulement de réussir, mais de savoir pourquoi on réussit.

Partagée, l'évaluation est moins descendante, n'apparaît plus comme sanction mais comme délibération porteuse de sens. Les élèves questionnent le niveau de maîtrise, hésitent, réalisant que l'enseignant procède de même. Se construit un système de normes linguistiques et langagières à partir de la grille commune, dont l'apprenant s'empare pour mesurer la valeur de son travail au regard de celui du groupe. Chaque élève peut aussi évaluer sa propre copie, sans plus de sévérité ni d'indulgence, pour progresser individuellement et mieux aider le reste de la classe.

Le statut de l'erreur évolue : les élèves voient celles des autres et relativisent les leurs. Si la circulation de copies est parfois conflictuelle, l'anonymisation permet ici d'éviter l'écueil d'un jugement biaisé par des pressions socioaffectives. La dématérialisation par l'écran et la distribution aléatoire effacent l'identité des auteurs. On évalue des textes et non des personnes.

⁷²Docimologie : Science des examens et des concours, étude de la qualité et de la validité des différents systèmes de notation scolaire et de contrôle des connaissances (source : cnrtl.fr).

Faire groupe avec les élèves

L'enseignant peut aussi faire groupe *avec* ses élèves, notamment en co-construisant la grille de coévaluation, ce qui entérine dans les esprits la distinction entre les consignes et les attendus qu'elles impliquent : un fructueux travail de reformulation émerge. En fonction des besoins, des envies, des contraintes formelles, structurelles, grammaticales ou stylistiques sont ajoutées, qui ne sont plus vécues comme des limitations mais comme des moyens d'obtenir un texte plus conforme, plus efficace ou plus beau.

Lors de l'évaluation de leurs pairs, les élèves ne travaillent plus simplement pour obtenir une bonne « note de coévaluateur⁷³ ». Ils savent que d'autres lisent leur texte, le prennent en compte. Leurs commentaires auront un impact bien réel sur les scripteurs. L'entraide dont le groupe est capable décentre l'enseignant de son rôle de seul référent, et lui offre la possibilité de participer aux exercices, au lieu de seulement les superviser, prêchant par l'exemple plus que par des modèles à imiter et bénéficiant, sans faux-semblants ni démagogie, de l'aide des élèves. Les écueils comme les potentialités des consignes initialement fournies se font jour lorsqu'on les met soi-même en pratique.

Chacun à sa mesure apporte sa pierre à l'édifice, si sa parole et ses réactions, ses doléances, remarques et conseils construisent la suite du projet de la classe.

Faire groupe avec les collègues

En troisième lieu, une nouvelle manière de faire équipe avec des collègues s'est construite. Nous avons l'habitude de collaborer avec nos pairs, dans divers espaces d'échanges, impulsés par les IA-IPR ou dans le cadre de la mise en œuvre de classes coopératives : cela nous permet de diversifier nos pratiques pour un enseignement plus efficace, favorisant la coopération et l'inclusion.

Le dispositif de coévaluation a déclenché un travail collaboratif distancié et présentiel entre des professeurs et un chercheur, avec des enjeux nouveaux : proposer une tâche pertinente pour le projet de recherche et pour la progression de la classe, évaluable par les élèves, avec des critères limités en nombre et transparents, ce qui résout une part des dilemmes que pose l'évaluation.

Face aux mêmes questionnements, le chercheur a partagé avec chaque participante les réponses que les autres avaient apportées, constituant un groupe de réflexion sur nos pratiques évaluatives.

Quelles perspectives pour la suite ?

En classe

Il serait intéressant d'utiliser le dispositif plus régulièrement, afin d'automatiser la procédure pour limiter la perte de temps « technique », d'inciter les élèves à passer toute

⁷³Moodle Atelier génère automatiquement deux notes par élève : une note de scripteur et une note de coévaluateur. Celle-ci mesure la précision avec laquelle un élève a évalué les copies de ses pairs et/ou sa propre copie, par comparaison avec l'évaluation de l'enseignant, qui constitue en quelque sorte « l'étalon de mesure » du dispositif.

production au crible d'attentes qui deviendraient celles du groupe et de l'élève. L'atelier Moodle créé est duplicable, avec des critères généraux prédéfinis, personnalisables : des activités de production écrite coévaluée pourraient ainsi être mises en œuvre facilement et régulièrement.

En formation

Nous souhaiterions partager cette expérience pour diffuser l'usage de Moodle Atelier et poursuivre la réflexion collective sur l'évaluation – notamment par compétences. L'appropriation par le groupe classe des normes linguistiques et langagières, rendue possible par la maîtrise du paramétrage de l'atelier, fait de ce dernier un objectif de formation.

Les fruits de cette coexpérimentation bénéficieraient également à des enseignants d'autres disciplines. Les groupes multidisciplinaires formés par Pierre Cieutat aux pédagogies coopératives ont justement pour habitude de mettre en commun des pratiques, dont les élèves peuvent être les porteparoles.

Dans un des collèges partenaires, sous l'impulsion de la cheffe d'établissement, un temps d'échange a été mis en place avec des collègues de disciplines diverses, qui se sont pliés au jeu de l'écriture puis de sa coévaluation : ce test a permis des échanges fructueux qui ont fait écho à ceux menés sur l'évaluation par compétences et sans notes. Ainsi, le travail collectif se poursuit.

Pour la recherche

D'après nos premiers résultats, un nombre significatif d'élèves auraient progressé en termes de compétences, d'autonomie et de coopération suite à cette coexpérimentation pédagonumérique menée grâce au concours des CARDIE et DRANE de Paris et Montpellier. Ce projet souligne l'intérêt du travail en groupe à trois niveaux : pour les élèves entre eux, pour l'enseignant avec ses élèves, mais aussi entre enseignants, chercheurs et partenaires des rectorats d'académie. Les connaissances scientifiques ainsi produites seraient peut-être d'autant plus pertinentes qu'elles seraient co-construites avec l'ensemble de la communauté éducative.

TABLE DES ILLUSTRATIONS

Figure 1 – Pratiques déclarées de différenciation pédagogique par les enseignants des pays de l'OCDE : données de l'enquête internationale TALIS 2013 (OCDE, 2014), reprises par (Cnesco, 2017).....	28
Figure 2 – Modélisation de la DP par (Forget & Lehraus, 2015, p. 75).....	31
Figure 3 – Coopération versus collaboration selon (Connac, 2017b, p. 20).....	36
Figure 4 – Continuum de l'autodétermination selon (Ryan & Deci, 2000) rapporté par (Carré, 2010).....	46
Figure 5 : La genèse instrumentale, combinaison de deux processus (Trouche, 2007).....	66
Figure 6 : Les niveaux d'emboîtement des instruments (Marquet, 2005a).....	67
Figure 7 – Exemple de « tableau de bord » d'un Atelier Moodle	70
Figure 8 – Schème d'utilisation "inter-évaluation en double aveugle"	72
Figure 9 – Attribuer le rôle "Etudiant" à l'enseignant.....	72
Figure 10 – Attribuer la pondération 16 à l'hétéro-évaluation de l'enseignant "déguisé" en "Etudiant"	73
Figure 11 – Formulaire d'évaluation avec échelle de Likert à cinq points	74
Figure 12 – Limitation à 4 coévaluations par élève lors de l'attribution aléatoire des travaux.....	75
Figure 13 – Modèle parallèle et convergent d'après (Creswell & Creswell, 2017)	82
Figure 14 – Modèle séquentiel explicatif.....	83
Figure 15 – Modèle séquentiel exploratoire-.....	83
Tableau 1 – Matrice de la recherche-intervention (Marcel, 2016a, p. 227).....	85
Figure 16 – Modèle d'intégration par combinaison	98
Figure 17 – Modèle parallèle et convergent d'après (Creswell & Creswell, 2017, p. 210).....	98
Figure 18 – Représentation schématique du protocole expérimental	101
Figure 19 – Exemple d'échelle de Likert à quatre points pour la coévaluation dans Moodle Atelier.....	103
Figure 20 – Critères de coévaluation pour la séance T1/T2 avec la classe A (3e)	104

Figure 21 – Critères de coévaluation pour la séance T1/T2 avec la classe A' (4e)....	105
Figure 22 – Critères de coévaluation à T1/T2 pour la classe B (3e).....	106
Figure 23 – Critères de coévaluation à T1/T2 pour la classe C (5e).....	107
Figure 24 – Critères de coévaluation à T3/T4 pour la classe A (3e)	107
Figure 25 – Critères de coévaluation à T3/T4 pour la classe A' (4e) : 1 ^{ère} question .	108
Figure 26 – Critères de coévaluation à T3/T4 pour la classe A' (4e) : 2 nd e question	108
Figure 27 – Critères de coévaluation à T3/T4 pour la classe B (3e).....	109
Figure 28 – Critères de coévaluation à T3/T4 pour la classe C (5e).....	110
Figure 29 – Critères de coévaluation à T5/T6 pour la classe C (5e).....	111
Figure 30 – Grille de coévaluation pour la classe B' (3 ^e , année N+1, 1 ^{ère} itération hors protocole)	113
Figure 31 – Grille de coévaluation pour la classe B' (3 ^e , année N+1, 2 nd e itération hors protocole)	114
Tableau 2 – Opérations d'instrumentalisation pour la catachrèse "coévaluation" du protocole semi-expérimental.....	116
Figure 32 – Exemple de grille de coévaluation avec écrit commentatif entre élèves	117
Figure 33 – Attribution des travaux dans Moodle Atelier : modes manuel et aléatoire	117
Figure 34 – Représentation chronologique de la recherche, avec temps de mesure et échantillons concernés	118
Tableau 3 – Structure détaillée du questionnaire en ligne en fonction des temps de mesure	121
Figure 35 – Exemple d'item avec échelle de Likert et descriptif (interface participants)	123
Figure 36 – Exemple d'item de calibrage.....	124
Figure 37 – Item unique mesurant le SEC-IAP dans le questionnaire en ligne.....	127
Figure 38 – Question fermée (item 1) portant sur la genèse instrumentale du dispositif expérimental.....	128
Figure 39 – Question ouverte (item 2) portant sur la genèse instrumentale du dispositif expérimental.....	129
Figure 40 – Question libre (item 3) portant sur la genèse instrumentale du dispositif expérimental.....	130
Tableau 4 – Les trois items visant à mesurer l'ergonomie du questionnaire selon les participants.....	130

Figure 41 – Exemple avec échelle de Likert à cinq points et descriptif	131
Tableau 5 – Grille indicatrice pour l'interprétation de la valeur de p, d'après (Singh, 2013)	138
Tableau 6 – Résultats du test de Shapiro-Wilk pour l'échelle SEPE	139
Tableau 7 – Conditions de déroulement des entretiens semidirigés et du focus groupe avec les enseignantes partenaires	149
Figure 42 – Interface utilisateur dans Tripetto® : éléments de ludification	151
Figure 43 – Accessibilité du questionnaire dans Moodle	151
Figure 44 – Dédoublage du questionnaire pour pallier les surcharges sur le serveur du collège	152
Tableau 8 – Hypothèses opérationnelles, données afférentes et analyses pratiquées	156
Figure 45 – Augmentation modérée des scores individuels de scripteurs de SE entre T2 et T4	158
Figure 46 – Augmentation du score de scripteur moyen de C sur T2-T6 (scores moyens exprimés en %, N = 22)	158
Figure 47 – Augmentation modérée des scores individuels moyens des élèves de SE à la sous-échelle SEP(scol) entre T3 et T4	159
Figure 48 – Forte augmentation des scores individuels moyens des élèves de A' à la sous-échelle SEP(scol) sur T3-T4	159
Figure 49 – Très forte augmentation des scores individuels moyens des élèves de C à la sous-échelle SEP(scol) sur T3-T4	160
Tableau 9 – Synthèse des résultats obtenus pour l'hypothèse HO 1 : Amélioration des compétences en français	161
Figure 50 – Augmentation modérée des scores individuels médians des élèves de SE à la sous-échelle SEP(srl) sur T1-T2	162
Figure 51 – Légère augmentation des scores individuels médians des élèves de SE à la sous-échelle SEP(srl) sur T3-T4	162
Figure 52 – Forte augmentation des scores individuels médians et moyens de B à la sous-échelle SEP(srl) sur T1-T2	163
Tableau 10 – Synthèse des résultats obtenus pour l'hypothèse HO 2a : amélioration de l'autorégulation des apprentissages	163
Figure 53 – Tendence haussière du score moyen de coévaluateur pour le super-échantillon SE (N = 61) sur T2-T4	165

Figure 54 – Hausse du score moyen de coévaluateur pour l'échantillon C (N = 23) sur l'intervalle T2-T6.....	165
Tableau 11 – Synthèse des résultats obtenus pour l'hypothèse HO 2b : amélioration de la métacognition.....	167
Figure 55 – Augmentation modérée des scores individuels moyens des élèves de SE à l'échelle SEP sur T1-T2.....	168
Figure 56 – Augmentation modérée des scores individuels moyens des élèves de SE à l'échelle SEP sur T3-T4.....	169
Figure 57 – Augmentation modérée des scores individuels moyens des élèves de A' à l'échelle SEP sur T3-T4.....	169
Figure 58 – Augmentation modérée des scores individuels moyens des élèves de C à l'échelle SEP sur T3-T4.....	169
Figure 59 – Forte augmentation des scores individuels moyens des élèves de B à l'échelle SEP sur T1-T2.....	170
Tableau 12 – Synthèse des résultats obtenus pour l'hypothèse HO 2c : amélioration du sentiment d'efficacité personnelle	170
Tableau 13 – Synthèse des résultats obtenus pour l'hypothèse HO 2 : Amélioration de l'autodirection des apprentissages	173
Figure 60 – Forte augmentation des scores SEP(coop) individuels moyens de A' sur T1-T2 (N = 8)	174
Figure 61 – Forte augmentation des scores SEP(coop) individuels moyens de B sur T1-T2 (N = 10)	174
Figure 62 – Augmentation modérée du score SEP(coop) individuel moyen de SE sur T1-T2 (N = 46)	175
Figure 63 – Très forte augmentation du SEC-IAP de A sur T3-T4 (N = 7).....	176
Figure 64 – Forte augmentation du SEC-IAP de SE sur T3-T4 (N = 30)	176
Tableau 14 – Synthèse des résultats obtenus pour l'hypothèse HO 3 : amélioration des apprentissages coopératifs	177
Tableau 15 – Combinaisons de résultats établissant la hausse du SEP(coop) de SE sur T1-T4	179
Figure 65 – Nuage de mots pour le corpus global « Retours écrits sur expérience », classe C, T6 (N = 23).....	185
Figure 66 – Nuage de mots pour le sous-corpus « Retours positifs », classe C, T6 (N = 23).....	185

Figure 67 – Nuage de mots pour le sous-corpus « Retours négatifs », classe C, T6 (N = 23)	186
Figure 68 – Graphe de similitude pour le sous-corpus « Retours positifs »	187
Figure 69 – Graphe de similitude pour le sous-corpus « Retours négatifs »	188
Tableau 16 – Conflits instrumentaux observés et artéfacts afférents (violet = artéfact pédagogique ; orange = artéfact technique ; bleu = artéfact didactique)	192
Tableau 17 – Conflits instrumentaux observés, avec artéfacts et degrés de conflictualité afférents (jaune = retard ; orange = renoncement partiel ; rouge = renoncement total)	194
Figure 70 – Rôle pivot de l'écrit commentatif dans le dispositif de coévaluation numérique	196
Tableau 18 – Conflits instrumentaux et synergies régulatrices (violet = artéfact pédagogique ; orange = artéfact technique ; bleu = artéfact didactique)	197
Tableau 19 – Niveaux de maîtrise et descriptifs issus de la réflexion docimologique menée par l'enseignante C avec ses élèves de 5 ^e à T1-T2 pour l'item « Comparaison » de la grille de coévaluation afférente.....	201
Figure 71 – Grille de coévaluation pour l'activité de coévaluation à T1-T2 avec l'échantillon A'	204
Tableau 20 – Principaux conflits instrumentaux observés et artéfacts afférents, du point de vue des enseignantes partenaires.....	219
Tableau 21 – Légende	219
Tableau 22 – Régulations instrumentales observées, avec artéfacts conflictuels et artéfacts régulateurs afférents	220
Tableau 23 – Evolution des schèmes d'utilisation et artéfacts afférents	221
Tableau 24 – Récapitulatif des quatre hypothèses opérationnelles et de leur validation	226
Tableau 25 – Contresens pédagogique et empêchement didactique : verbatim d'élèves et scores moyens de scripteurs et de coévaluateurs (T2-T4, en %)	228
Tableau 26 – Scores de scripteurs et de coévaluateurs des élèves du 3e quartile à T2 (par score de scripteur croissant, en %)	230
Tableau 27 – Scores de scripteurs et de coévaluateurs des élèves du 3e quartile à T4 (par score de scripteur croissant, en %)	231
Tableau 28 – Problèmes de discernement docimologique : verbatim d'élèves et scores afférents (en %).....	233

Figure 72 – Mésusages ou non-usages des interfaces associées à l’artéfact numérique : combinatoire conflictuelle	235
Figure 73 – Evolution du score de scripteur de Marie de T2 à T6 (en %).....	236
Figure 74 – Evolution du score de coévaluatrice de Marie de T2 à T6 (en %)	237
Figure 75 – Evolution des scores de scripteurs entre T2 et T4 pour l’échantillon B, après appariement (N = 11).....	241
Figure 76 – Evolution des scores de coévaluateurs entre T2 et T4 pour l’échantillon B, après appariement (N = 11).....	241
Figure 77 – Matrice de corrélation représentant les résultats du test de Spearman pratiqué sur les scores de coévaluateurs et les scores individuels moyens à l’échelle SEPE, échantillon SE, T2, N = 29 (en %, après appariement).....	249
Tableau 29 – Engagement (méta-)cognitif dans la formulation d’écrits commentatifs argumentés à destination des pairs coévalués	251
Tableau 30 – Verbatim d’élève illustrant la vision conflictuelle entre hétéro- et inter-évaluation.....	253
Tableau 31 –Verbatim d’élèves illustrant le souci d’aider le pair évalué avec un écrit commentatif adapté.....	257
Tableau 32 – Principaux conflits instrumentaux identifiés et artéfacts afférents, du point de vue des élèves.....	262
Tableau 33 – Légende	262
Figure 78 – Piste interprétative des interactions à l’œuvre dans la genèse instrumentale du dispositif.....	266
Tableau 34 – Principaux conflits instrumentaux observés et degrés de conflictualité afférents (point de vue élèves)	267
Tableau 35 – Légende	267
Tableau 36 – Principales régulations instrumentales observées du point de vue des enseignantes partenaires, avec artéfacts conflictuels, artéfacts affectés et artéfacts régulateurs afférents.....	269
Tableau 37 – Légende	270
Tableau 38 – Augmentation du score de coévaluateur pour SE (N = 61) entre T2 et T4	275
Figure 79 – Augmentation des score de coévaluateur pour SE (N = 61) entre T2 et T4	275
Figure 80 – Evolution du score de scripteur d’Etienne à T2, T4 et T6 (en %).....	278

Figure 81 – Evolution du score d'évaluateur d'Etienne à T2, T4 et T6 (en %)	279
Figure 82 – Evolution du score individuel médian d'Etienne à l'échelle SEPE à T1, T2, T3, T5 et T6 (sur échelle de 1 à 5).....	280
Figure 83 – Evolution du score individuel médian d'Etienne à l'échelle SEP(srl) à T1, T2, T5 et T6 (sur échelle de 1 à 5).....	280
Figure 84 – Augmentation des scores de scripteurs de SE (N = 58) entre T2 et T4..	281
Tableau 39 – Notes de scripteurs de SE (N = 58) à T2 et T4 : statistiques descriptives	282
Figure 85 – Médiane et écart interquartile pour les notes de scripteurs de SE (N = 58) à T2 et T4.....	282
Figure 86 – Evolution du score de scripteur moyen de C (N = 22) à T2, T4 et T6...	283
Tableau 40 – Notes de scripteurs de C (N = 22) à T2 et T4 : statistiques descriptives	283
Figure 87 – Médiane et écart interquartile pour les notes de scripteurs de C (N = 22) à T2, T4 et T6	283
Figure 88 – Evolution du score de coévaluateur moyen de SE (N = 61) à T2 et T4.	284
Tableau 41 – Notes de coévaluateurs de SE (N = 61) à T2 et T4 : statistiques descriptives	284
Tableau 42 – Notes de coévaluateurs de C (N = 23) à T2, T4 et T6 : statistiques descriptives	285
Figure 89 – Evolution des scores moyens de coévaluateurs de C (N = 22) à T2, T4 et T6.....	285
Figure 90 – Médiane et écart interquartile pour les notes de coévaluateurs de C (N = 23) à T2, T4 et T6	286
Figure 91 – Proposition de représentation schématique pour la catachrèse « Boucle de rétroaction formative anonymisée ».....	306
Figure 92 – Résultats des deux items relatifs à l'expérience préalable des élèves participants en matière d'auto- et d'inter-évaluation (valeurs moyennes à T1, échantillons A, A', B, C et SE, N = 62).....	337
Tableau 43 – Statistiques descriptives pour les 3 questions relatives à l'ergonomie du questionnaire	337
Figure 93 – Exemple de question d'entraînement.....	338
Figure 94 – Résultat du test de calibrage à T1	338
Figure 95 – Résultat du test de calibrage à T3	339

Tableau 44 – Tests de fiabilité pour l'échelle SEPE avec SE à T1, T2, T3 et T4 (N = 282)	339
Tableau 45 – Tests de cohérence interne pour l'échelle SEP(scol) avec SE à T1, T2, T3 et T4 (N = 282)	340
Tableau 46 – Tests de cohérence interne pour l'échelle SEP(srl) avec SE à T1, T2, T3 et T4 (N = 282)	340
Tableau 47 – Tests de cohérence interne pour l'échelle SEP(coop) avec SE à T1, T2, T3 et T4 (N = 282)	340
Figure 96 – Cohérence interne pour l'échelle SEPE et les sous-échelles SEP(scol), SEP(srl) et SEP(coop) avec SE à T1, T2, T3 et T4 (N = 282).....	340
Figure 97 – Item 1 : Analyse descriptive des réponses des élèves à T2 (N = 59) et T4 (N= 57).....	360
Figure 98 – Item 2 : principales thématiques à T2 (N = 59) et T4 (N = 57).....	361
Figure 99 – Item 3 : principales thématiques à T2 (N = 27).....	362
Tableau 48 – Verbatim d'élèves : présence de prérequis didactiques	362
Tableau 49 – Verbatim d'élèves : absence de prérequis didactique comme source de conflit instrumental	363
Tableau 50 – Verbatim d'élèves : difficulté à évaluer la production écrite d'un pair (discernement docimologique).....	363
Tableau 51 – Verbatim d'élève : qualité linguistique insuffisante de certains écrits et difficulté afférente à coévaluer	364
Tableau 52 – Verbatim d'élèves : appétence pour la coévaluation en soi.....	364
Tableau 53 – Verbatim d'élèves : difficulté à évaluer la production écrite d'un pair (peur de commettre une injustice)	365
Tableau 54 – Verbatim d'élèves : difficulté à évaluer la production écrite d'un pair (inter-éval vs. hétéro-éval)	365
Tableau 55 – Verbatim d'élèves : difficulté à formuler un écrit commentatif adapté	365
Tableau 56 – Verbatim d'élèves : importance de l'expérience acquise	366
Tableau 57 – Verbatim d'élèves : contresens concernant l'artéfact pédagogique	366
Tableau 58 – Verbatim d'élèves : facilité d'utilisation	367
Tableau 59 – Verbatim d'élèves : appétence pour l'informatique en soi	367
Tableau 60 – Verbatim d'élève : mésusages ou non-usages des interfaces associées à l'artéfact numérique	367
Tableau 61 – Verbatim d'élèves : artéfact numérique et étayage pédagogique	367

Tableau 62 – Verbatim d'élèves : anonymat	368
Tableau 63 – Verbatim d'élèves : apports de l'artéfact numérique à la compréhension des consignes	368
Tableau 64 – Scores moyens de scripteurs et de coévaluateurs des élèves ayant souligné l'importance de l'étayage	369
Figure 100 – Grille de coévaluation pour l'activité de production écrite / étude de la langue menée avec la classe C (5 ^e) à T1-T2	371
Figure 101 – « Windflowers » de John William Waterhouse (1903) : image support pour la classe C à T3/T4	372
Figure 102 – Texte support pour la séance T5/T6 avec la classe C (5e) : Le Sumo qui ne pouvait pas grossir, Eric-Emmanuel Schmitt (Albin Michel, 2009).....	373
Figure 103 – Aide à l'écriture (séance T5/T6, classe C, 5e)	373
Figure 104 – Formulaire d'évaluation pour la 1 ^{ère} réappropriation par l'enseignante B (classe B', 3 ^e , année N+1).....	374
Figure 105 – Formulaire d'évaluation pour la 2 ^{nde} réappropriation par l'enseignante B (classe B', 3 ^e , année N+1).....	375
Tableau 65 – Les 7 items de la sous-échelle SEP(scol).....	386
Tableau 66 – Les 10 items de la sous-échelle SEP(srI)	386
Tableau 67 – Les 5 items de la sous-échelle SEP(coop)	387
Figure 106 – Echelle de Likert à cinq points avec captions explicatives	388
Figure 107 – Information RGPD	388
Figure 108 – Exemple de ludification dans Tripetto®	389