



HAL
open science

Classifications psychiatriques : des catégories aux symptômes

Christophe Gauld

► **To cite this version:**

Christophe Gauld. Classifications psychiatriques : des catégories aux symptômes. Philosophie. Université Panthéon-Sorbonne - Paris I, 2023. Français. NNT : 2023PA01H212 . tel-04758605

HAL Id: tel-04758605

<https://theses.hal.science/tel-04758605v1>

Submitted on 29 Oct 2024

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

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

UNIVERSITÉ PARIS 1 PANTHÉON-SORBONNE & UNIVERSITÉ LYON 3
École Doctorale de Philosophie (Université Paris 1 – ED 280)
Institut d’Histoire et de Philosophie des Sciences et des Techniques (UMR 8590)
& Institut de Recherches Philosophiques de Lyon

THÈSE

Pour l’obtention du titre de Docteur en Philosophie
Présentée et soutenue publiquement
Le 20/12/2023 par
Christophe GAULD

Classifications psychiatriques :
Des catégories aux symptômes

Sous la direction :

Du Professeur Denis Forest (Université Paris 1 Panthéon-Sorbonne)
Et de la Professeure Élodie Giroux (Université Lyon 3)

Membres du jury :

M. Luc Faucher	Professeur à l’Université du Québec à Montréal	– Rapporteur
M. Denis Forest	Professeur à l’Université Paris 1	– Directeur
M. Pierre Fourneret	Professeur à l’Université Lyon 1	– Examinateur
Mme. Élodie Giroux	Professeure à l’Université Lyon 3	– Directrice
Mme. Lara Keuck	Professeure à l’Université de Bielefeld	– Rapporteur
M. Maël Lemoine	Professeur à l’Université de Bordeaux	– Examinateur

**Classifications psychiatriques :
Des catégories aux symptômes**

Remerciements

Je tiens à exprimer ma profonde gratitude envers le **Professeur Denis Forest**. Je suis reconnaissant pour la confiance qu'il a placée en moi, pour la considération qu'il a toujours manifestée envers mon travail et pour l'inspiration qu'il m'a fournie pour orienter ma thèse vers une réflexion argumentée autour de la pratique clinique.

Je souhaite également exprimer mes remerciements sincères à la **Professeure Élodie Giroux**. Ses conseils, ses enseignements et son soutien inébranlable au fil de ces années ont été d'une valeur inestimable. J'espère qu'elle ne sous-estime pas l'influence qu'elle a eu, tant sur le plan philosophique qu'humain, sur mon parcours professionnel et mon parcours de vie plus généralement.

Mes remerciements s'adressent également à la **Professeure Lara Keuck** et au **Professeur Luc Faucher** pour le temps qu'ils ont consacré à examiner ce travail. La lecture de leurs travaux a été particulièrement précieuse pour ma thèse et a influencé ma réflexion de manière plus générale.

Je tiens à exprimer ma gratitude envers le **Professeur Pierre Fourneret** et le **Professeur Maël Lemoine**, qui ont accepté de faire partie de mon jury. Mon parcours clinique m'a permis de croiser le chemin du **Professeur Pierre Fourneret**, qui a toujours reconnu l'importance des travaux philosophiques pour notre pratique clinique quotidienne. Sa confiance a été inestimable. De même, mon parcours philosophique m'a conduit à découvrir la pensée du Professeur Maël Lemoine, qui m'a ouvert les portes à cette riche perspective des sciences médicales.

Je tiens à souligner que ce travail n'aurait pas été possible sans le soutien précieux de nombreuses personnes dans ma vie. Je remercie tout particulièrement **Agathe**, **Basile**, mes **parents**, mes **beaux-parents**, mes **frères**, ma **grand-mère** et mes **amis**. Leur présence constante a été un véritable soutien et je leur suis reconnaissant d'avoir toujours été là pour moi.

Sommaire

INTRODUCTION GÉNÉRALE

LES MULTIPLES FONCTIONS DU MANUEL DIAGNOSTIQUE ET STATISTIQUE DES TROUBLES MENTAUX : UN REFLET DE SON HISTOIRE

- I. Avant le DSM
 - II. Les critères de Feighner et les critères diagnostiques pour la recherche
 - III. L'élaboration du DSM-III
 - IV. L'évolution des versions du DSM après le DSM-III
- Conclusion

LE PROJET RESEARCH DOMAIN CRITERIA : DES NEUROSCIENCES À L'ENVIRONNEMENT

- Introduction
- I. Description de la matrice des RDoC
 - II. Genèse des RDoC
 - III. Validité de construit des RDoC
 - IV. Les RDoC et les trois niveaux d'intégration
 - V. RDoC, intégration du neurodéveloppement et application clinique
- Conclusion

LA TAXONOMIE HIÉRARCHIQUE DE LA PSYCHOPATHOLOGIE : À LA CROISÉE ENTRE CLINIQUE ET STATISTIQUES

- Introduction
- I. Psychologie dimensionnelle, nosologie quantitative et statistiques
 - II. La structure dimensionnelle de la HiTOP
 - III. La structure hiérarchique de la HiTOP
 - IV. Discussion des enjeux
- Conclusion

TRANSITION

DES APPROCHES PAR LA CATÉGORIE ET DIMENSION AUX APPROCHES PAR LE SYMPTÔME

- Introduction
- Postulat I – Dysfonction d'un trouble psychiatrique et mécanismes locaux
- Postulat II – Préjudice causé par une dysfonction et analyse du symptôme préjudiciable
- Une définition potentiellement pertinente de ce qu'est un trouble psychiatrique pour la pratique clinique
- Conclusion

CONCLUSION GÉNÉRALE : DES CATÉGORIES AUX SYMPTÔMES

Avertissements

1. Les traductions de texte en langue étrangère sont les miennes.
2. Lorsque le vocabulaire scientifique a été proposé dans une langue étrangère, j'ai essayé de proposer dans la mesure du possible une traduction française. Cependant, lorsque cette traduction n'est pas communément admise et employée, j'ai préféré conserver le terme (anglosaxon) original, alors indiqué en italique. Par exemple, pour le terme « *hub* » épistémique, j'ai proposé la traduction de « plaque tournante » épistémique mais dans la suite du texte, je parle de *hub* épistémique.
3. La liste des abréviations employées est fournie ci-après.
4. J'ai fait le choix de limiter les notes de bas de page pour alimenter des annexes, elles-mêmes les plus concises possibles. Les annexes ne sont pas nécessaires à la compréhension du fil argumentatif de la thèse, mais restent importantes pour la compréhension générale du propos. Je propose au lecteur de consulter de manière ciblée ces annexes lorsqu'elles sont appelées dans le texte sous le format (**Annexe**, « **Titre** »). De plus, un document d'annexes originales, mises en forme (635 pages réparties en 9 annexes), complémentaires mais non nécessaires à ce travail, peut m'être demandé expressément.
5. Dans la suite de ce travail, j'ai adopté l'usage du « nous » de courtoisie.

Table des figures

Figure 1 – Synthèse introductive de l’organisation des chapitres de la thèse	32
Figure 2 – La matrice des <i>Research Domain Criteria</i> (RDoC)	83
Figure 3 – Localisation des fonctions corticales selon l’architecture cérébrale	87
Figure 4 – Les construits diagnostics du DSM-5 correspondant aux catégories DSM des troubles de personnalité mis en relation avec les principales fonctions cérébrales implémentées dans les <i>Research Domain Criteria</i>	115
Figure 5 – Modèle du consortium de travail de la Taxonomie Hiérarchique de la Psychopathologie (<i>Hierarchical Taxonomy of Psychopathology</i> , HiTOP)	135
Figure 6 – Seuil de préjudice considéré comme pathologique selon les valeurs de la société	197

Tableaux

Tableau 1 – Caractéristiques de l’approche par la catégorie / dimension.

Tableau 2 – Équivalence des termes renvoyant à l’approche par la catégorie / dimension et à l’approche par la symptomé.

Liste des abréviations

APA : *American Psychiatric Association*

CIM : Classification internationale des maladies

DSM : *Diagnostic and Statistical Manual of Mental Disorders*

FDA : *Federal Drug Administration*

HDA : *Harmful Dysfunction Analysis*

HiTOP : *Hierarchical Taxonomy Of Psychopathology*

IRT : *Item Response Theory*

NIMH : *National Institute of Mental Health*

OMS : Organisation Mondiale de la Santé

QI : *Quotient Intellectuel*

RDC : *Research Diagnostic Criteria*

RDoC : *Research Domain Criteria*

ROAMER : *Roadmap Mental Health Europe*

INTRODUCTION GÉNÉRALE

Une jeune patiente de 13 ans est accueillie en hospitalisation pour des épisodes de dissociation intenses associés à des hallucinations acoustico-verbales. Elle présente également un mutisme sélectif, masquant une grande pertinence dans l'analyse des situations sociales. Elle nous parle elle-même de « système-hôte », terme renvoyant à une forme de cyberchondrie associée à une condition décrite comme un « trouble dissociatif de l'identité » auto-diagnostiqué. Aucun autre élément sémiologique n'est évident, hormis une intense détresse qu'elle exprime par un visage fermé et un repli. Bien qu'elle puisse décrire un profond sentiment de vide, elle ne décrit pas de symptomatologie franche pouvant précisément orienter vers un trouble psychiatrique. Il s'agit d'un cas relativement typique pouvant s'apparenter au « trouble dépressif caractérisé » tel qu'il peut se manifester en pratique clinique.

Louis a 12 ans. Il convulse cinq fois par jour depuis des mois. Ses parents doivent l'envoyer régulièrement aux urgences en fauteuil roulant car il ne parvient plus à marcher. Sa mère est porteuse d'une sclérose en plaque, son père d'une maladie métabolique invalidante et sa sœur une malformation cardiaque. Pour lui, aucun bilan paraclinique n'a décelé d'anomalie malgré la passation de lourdes batteries d'exams médicaux. Il n'a aucune plainte verbale, ne décrit pas de souffrance. Il peut sourire et se montrer avenant avec tous. Il s'agit d'un cas relativement typique pouvant s'apparenter au « trouble de conversion », ou trouble à symptomatologie neurologique fonctionnelle.

Adèle a 14 ans. Elle a été abandonnée au Congo par ses parents quand elle était petite. Ceux-ci, pris de remords de cet abandon, l'on fait venir en Angola où ils avaient déménagé. Les violences physiques et sexuelles de la part du père ont débuté à son arrivée. Elle a été envoyée secrètement en France par sa mère. Elle y a alors commis plusieurs tentatives de suicide, notamment par immolation ou par saut dans le vide. Elle raconte son histoire de manière désaffectivée, exprime des symptômes de tristesse qui s'expriment par de l'irritabilité et des crises de colère et de l'anxiété succédant aux épisodes de

reviviscences et aux cauchemars. Elle semble profondément affectée par son histoire de vie, décrivant une profonde souffrance morale – qu'elle ne parvient pas à décrire ni à verbaliser. Il s'agit d'un cas relativement typique pouvant s'apparenter au « trouble-stress post-traumatique ».

Ces tableaux cliniques, riches sur le plan sémiologique, sont parmi les plus classiques de la (pédo)psychiatrie. Quel que soit le niveau de pratique et d'expérience du clinicien, transformer ces situations en symptômes labellisés est un travail particulièrement délicat. Quand bien même ces situations parviendraient à être « traduites » dans un langage médical, pourrait-on vraiment espérer que chacune d'elle puisse être associée à une (ou même plusieurs) catégorie(s) psychiatrique(s) ? Dans ces situations, il n'est pas évident qu'une catégorie nosographique puisse s'appliquer à l'expérience vécue et décrite par ces patients. En fait, aucun de ces tableaux cliniques ne remplit les critères du Manuel diagnostique et statistique des troubles mentaux (*Diagnostic and Statistical Manual of Mental Disorders* – DSM) (American Psychiatric Association 2013). Il manque pour chacun un ou plusieurs critères qui permettraient de statuer sur le diagnostic tel qu'il est décrit dans cette classification. Le clinicien s'appuie cependant sur ces descriptions et symptômes pour orienter sa compréhension du tableau clinique et ses prises en charge, comme la prévention du suicide, le traitement médicamenteux ou l'information des familles et des travailleurs sociaux. Le contraste est saillant entre ces situations cliniques et la rigueur systématique des nosographies que nous présenterons dans la suite de ce travail. En pratique, face à cette dissonance, le clinicien a grossièrement deux options : *s'obstiner* à appliquer les critères nosologiques à son patient, quitte à en déformer les propos ou à céder à une interprétation subjective de ceux-ci ; ou bien *mettre temporairement ces critères de côté* pour recueillir les symptômes tels qu'il apparaissent dans l'entretien clinique, recueil qui pourra ensuite (éventuellement) être appliqué à telle ou telle catégorie nosographique. Parler même de diagnostic peut être problématique en pratique, ajoutant encore une tension entre conceptions et classifications et pratique clinique. Considérons par exemple une entité débattue dans la littérature et la pratique médicale, la dysphorie de genre d'apparition rapide (*Rapid Onset Gender Dysphoria* – ROGD) (Littman, 2018). Celle-ci correspond à l'apparition et à la disparition

relativement « rapide » d'une revendication de dysphorie de genre, parfois sur moins de trois jours. Une telle situation clinique est actuellement particulièrement répandue en pratique clinique pédopsychiatrique¹. Les débats (notamment épistémologiques) ont cherché à comprendre si cette espèce diagnostique (la ROGD) devrait ou non être incluse dans les recommandations de bonnes pratiques et / ou dans les classifications. Un nombre conséquent de recherches attestent que cette condition ne peut ni intégrer une classification, ni être force de propositions politiques et de santé publique. Il a donc été (socialement, scientifiquement, institutionnellement et politiquement) admis que l'apparition et la disparition d'une dysphorie de genre ne peut faire l'objet d'une standardisation au sein des classifications. Cependant, ce phénomène est bien présent en clinique : le clinicien ne peut nier qu'un ou une jeune revendique ressentir un mal-être (certes passager) lié à son identité². Il existe bien une souffrance décrite par un individu et cette souffrance s'exprime autour d'un idiome de détresse qu'est la question du genre. Ce qui intéresse le clinicien n'est pas de savoir s'il « existe » une ROGD en tant qu'entité individualisée. Il s'attache plutôt à l'ensemble de symptômes qui lui permettent d'agir en pratique.

Considérons encore un autre exemple permettant de mieux saisir les relations entre catégories et pratique clinique. Un homme de 70 ans est reçu au service médico-psychologique régional de la prison dans laquelle il vient d'être incarcéré. Le clinicien doit le recevoir en raison des idées suicidaires qui sont apparues plusieurs mois après son incarcération. Lors de la consultation, il ne présente plus aucun symptôme, mais lorsque le dossier est consulté, le clinicien remarque qu'un diagnostic a été posé, comme cela est nécessaire sur le plan légal pour toute consultation en milieu carcéral : « Trouble de l'adaptation », y lit-on. Ce diagnostic renvoie à une détresse marquée qui est hors de proportion par

¹ Notons que la dysphorie de genre que nous discutons ici est légitimement médicalisable car elle diffère de l'incongruence de genre. La dysphorie correspond à un mal-être décrit par la personne concernant son identité de genre. L'incongruence de genre renvoie à une condition qui questionne l'identité elle-même (et qui par définition ne devrait pas relever à elle seule de la médecine).

² Quoique cet exemple pourrait paraître différent d'autres cas rencontrés en psychiatrie clinique, parce qu'il touche à la question de l'identité, il faut bien préciser que tout trouble touche de près ou de loin à la question des normes d'un individu – et que, de fait, toute condition et tout trouble psychiatrique nécessite de considérer ces normes.

rapport à la sévérité ou à l'intensité du facteur de stress. Le diagnostic est utile sur le plan administratif, dans le cadre de l'incarcération, nécessaire au remboursement des consultations médicales. Après avoir éliminé d'autres facteurs traumatiques ou de stress, le clinicien pourrait se demander si le milieu carcéral déshumanisant impose de poser un tel diagnostic (diagnostic qui ne serait alors qu'une des conséquences d'un environnement pathogène). Le clinicien se retrouve ainsi dans une situation particulièrement désagréable, dans laquelle le diagnostic lui est « imposé » par un environnement (ici, pathogène) et des contraintes (ici, administratives) et son travail clinique serait alors celui de chercher une harmonisation entre des critères diagnostiques et les symptômes d'un tableau clinique. Cette nécessité d'harmonisation est une constante dans la pratique clinique. Le rôle du clinicien sera de tenter d'ajuster la catégorie à la symptomatologie (particulièrement pauvre) de son patient. Cependant, il pourrait également prendre le parti d'accepter que le diagnostic ne puisse avoir de fonction que sur le plan administratif. S'il choisit cette deuxième option (ce qui semble être le cas d'un grand nombre de cliniciens éclairés), une espèce psychiatrique n'est alors pas définie parce qu'elle ressemble au tableau clinique d'un patient, mais parce qu'elle lui permet d'obtenir un soin médical lui offrant un certain nombre de ressources dans un milieu préjudiciable. Dans cet exemple, le « trouble de l'adaptation » est une catégorie qui n'est qu'utile sur le plan administratif et ne semble pas pertinente pour guider la pratique clinique³.

En partant de ce genre de problématiques cliniques, nous souhaitons dans cette thèse nous questionner sur l'écart entre catégories et symptômes. Un certain nombre de recherches ont montré que l'utilisation des catégories et classifications dans la pratique clinique nécessitait une certaine souplesse. Cependant, renoncer à recourir (même temporairement) aux catégories et classifications lors de la consultation clinique nécessite tout de même d'ordonner et structurer la sémiologie clinique. Cette question n'est pas tant celle relevant d'un problème de raisonnement ou de théorie de la décision médicale, mais bien une question épistémologique : il s'agit de comprendre comment les catégories (des nosographies) et les

³ D'ailleurs, les résultats d'enquêtes sur le terrain soutiennent cette idée : un grand nombre d'utilisateurs du DSM juge que ses catégories ont une faible utilité pour « choisir un traitement » ou « évaluer le probable pronostic » (First et al., 2018).

symptômes (issus de la pratique clinique) se combinent et s'agencent⁴. La philosophie de la psychiatrie s'est d'ailleurs développée à son origine autour des questions de définition et de catégorisation des troubles psychiatriques. Au cours des dernières décennies, un certain nombre de recherches dans ce champ ont été justement consacrées à la question des espèces psychiatriques (Haslam 2002 ; Kendler et al. 2011 ; Zachar 2000, 2015). C'est cette notion d'espèce psychiatrique qui nous servira principalement à répondre à la question de la relation entre catégories et symptômes. Les débats sur les espèces psychiatriques sont de nature ontologiques ou épistémologiques. La question ontologique concerne ce qu'est, *en soi*, un trouble psychiatrique. Dans ce travail, nous adoptons plutôt une conception épistémologique des espèces psychiatriques. Nous nous intéressons aux espèces psychiatriques en tant qu'elles reflètent les conceptions et classifications adoptées par les cliniciens. Par conséquent, notre acception du terme d'espèce psychiatrique englobe des désignations comme celles de « trouble », « syndrome », « catégorie », « dimension », « maladie » ou « pathologie ». Nous constatons une invariance dans les débats sur les espèces psychiatriques : tous les types d'espèces psychiatriques (conçues comme espèces naturelles, espèces pratiques, espèces complexes, espèces floues ou espèces interactives – Haslam parlait « d'espèces d'espèces ») sont des « espèces diagnostiques ». Une espèce diagnostique est une espèce psychiatrique correspondant à une entité non-observable (ce à quoi renvoie la définition de « l'espèce ») constituée d'ensembles d'éléments observables (les éléments constituant l'espèce, par exemple les symptômes servant au diagnostic). En psychiatrie, parler d'espèce diagnostique revient à considérer les troubles psychiatriques en tant que catégories, dimensions ou spectres (par exemple, le « spectre de la schizophrénie »). De telles catégories, dimensions ou spectres ont été codifiées dans des taxonomies comme le DSM ou la Classification internationale des maladies (CIM)

⁴ L'une des réponses à ce genre de questions peut être retrouvée dans les théories de la prise de décision décrites en pédagogie médicale ; nous faisons cependant le choix de ne pas aborder ces théories, qui nécessiteraient de nous pencher sur la manière dont raisonne les cliniciens. Nous pensons qu'une étude de la structuration des symptômes, (relativement) indépendante des facteurs expliquant le raisonnement du clinicien, sera plus féconde.

(Cooper 2005; Organisation mondiale de la santé 2017)⁵. Dans cette thèse, nous nous intéresserons à la conception de l'espèce psychiatrique qu'a implicitement le clinicien dans sa pratique. Par pratique « clinique », nous désignons la pratique du clinicien *lors de l'entretien* ou la consultation médicale, face à son patient – renvoyant au *klinikein*, « au lit du malade » (et non pas sa pratique plus générale en tant que clinicien au sein d'une institution, qui requiert un codage administratif, une communication avec d'autres soignants, une présentation de ses cas lors de réunions, etc.).

L'approche par la catégorie et la dimension

La multiplicité des espèces psychiatriques n'est pas sans rapport avec la multiplicité des classifications psychiatriques. L'émergence des différentes espèces psychiatriques dépendrait de l'émergence des différentes classifications psychiatriques – émergence relativement récente dans l'histoire de la psychiatrie. Ainsi, il existerait plusieurs définitions de ce que serait une espèce psychiatrique en fonction de différentes classifications psychiatriques. On pourrait alors parler de pluralisme « définitionnel », et de pluralisme « nosographique », horizontaux et non-hiérarchiques. Des exemples d'un tel pluralisme nosographique sont les suivants : le projet neuroscientifique de l'Institut national de la santé mentale des *Research Domain Criteria* (RDoC), axé autour d'une matrice constituée d'éléments neuroscientifiques (Cuthbert 2014 ; Ferrante et al. 2019), dont les construits dimensionnels peuvent être considérés comme des espèces naturelles ; ou la *Hierarchical Taxonomy Of Psychopathology* (HiTOP), dont les clusters

⁵ Nous pouvons déjà mentionner ici que c'est en partie du fait de cet intérêt ciblé sur les espèces diagnostiques qu'il existe une profonde discordance entre les propositions actuelles en philosophie de la psychiatrie (cherchant par exemple à délimiter le normal du pathologique) et la perception des cliniciens (qui veillent avant tout à identifier et ordonner – hiérarchiser, structurer – les symptômes d'un patient). Pour une grande part, l'essentiel d'un entretien clinique est destiné à recueillir les entités élémentaires que présente un patient, allant de facteurs socio-démographiques à des émotions relativement complexes, sans toutefois que le travail de synthèse de ces différentes entités élémentaires, tel qu'il est effectué par le clinicien, ne nécessite un chevauchement parfait avec la catégorie diagnostique.

hiérarchiques peuvent être considérés comme des espèces statistiques – car regroupés selon les variables latentes issus d’analyses factorielles (Kotov et al. 2017). La compréhension de la multiplicité des espèces psychiatrique nécessite donc d’étudier la multiplicité de classifications.

Nous soutenons de plus que chacune de ces classifications est le témoin d’une conception de la psychiatrie. Exprimé autrement, chaque classification est la cristallisation d’une certaine conception de la psychiatrie. Ce dernier point est important pour comprendre notre intérêt pour les classifications : afin de ne pas avoir à discuter les différentes conceptions de la psychiatrie – la notion de conception étant particulièrement large – nous discutons des classifications de la psychiatrie. Plus précisément, l’un des principal écueil dans lequel nous souhaiterions ne pas tomber pour ce travail serait celui de se lancer dans un débat sur la nature de ces « conceptions », débat qui a hanté la nosologie en en faisant tantôt des styles de pensée, des styles de raisonnement, des programmes de recherche, des matrices disciplinaires, des répertoires, des paradigmes, etc. En défendant l’idée que chaque classification psychiatrique est une instanciation d’une conception de la psychiatrie, nous admettons que toute classification véhicule des méthodes, concepts, théories, terminologies et valeurs de différentes communautés cliniques et scientifiques de cette conception⁶. Dans ce travail, la notion de conception reste donc volontairement floue, mêlant des éléments sociologiques, communautaires, pratiques et philosophiques⁷. Il est d’ailleurs plus aisé d’étudier les définitions des espèces psychiatriques au travers des classifications (en tant qu’instanciations de différentes conceptions de la psychiatrie) qu’au travers des conceptions de la

⁶ Ludwik Fleck a largement abordée la notion de communauté scientifique, en en décrivant les différents facteurs permettant de relier collectivement les membres d’une telle communauté (par exemple, des facteurs émotionnels, les pressions internes et externes, les « mystagogies » permettant l’intégration, etc.). Cependant, le recours aux travaux de Fleck nécessiterait d’étendre ses thèses à plus d’une seule communauté psychiatrique, ce que nous n’aborderons pas dans ce travail. En étudiant les classifications en tant qu’objets de médiation pratiques pour le clinicien et le philosophe, nous considérons qu’elles sont les mises en application (ou substantialisations) des conceptions de la psychiatrie – et donc de courants et traditions cliniques et scientifiques. Nous soutenons que les classifications offrent avant tout la possibilité de manier des outils concrets, permettant de discuter des manières de concevoir la psychiatrie. Elles sont donc considérées dans ce travail comme les témoins de l’appropriation et de la diffusion d’une compréhension du monde et non pas uniquement comme des structures cliniques ou scientifiques ordonnant des espèces (Bird & Tobin, 2008).

⁷ Si la notion de « conception » fait certes penser à celle de paradigme ou de style de pensée, nous choisissons de conserver un terme plus neutre pour ne pas nous référer (et justifier l’usage) de telle ou telle philosophie. À noter que cette thèse ne cherche donc pas à étudier la nature de la psychiatrie elle-même, par exemple en tant qu’elle pourrait être considérée comme une science ou comme un art.

psychiatrie. Par exemple, dans le cadre des troubles du neurodéveloppement (comme le trouble du spectre de l'autisme (TSA) ou le trouble déficit de l'attention avec ou sans hyperactivité (TDAH)), certains cliniciens développeront une forme de « goût » (un habitus, un style) pour la neurobiologie et ils fonderont leurs pratiques sur l'usage de prélèvements sanguins, de réunions de concertation pluridisciplinaires avec les équipes de génétique, de consultations « expertes », d'inclusions dans les cohorte destinées à déceler des marqueurs morphologiques ou à utiliser des modèles (mathématiques) de croyances et de prédictions pour leurs entretiens cliniques. Leur prise en charge du TSA ou du TDAH sera orientée selon une feuille de route de lignes médicamenteuses guidée par les récepteurs spécifiques aux neurotransmetteurs impliqués dans tel ou tel trouble. Ils adoptent donc une conception plus biologique de la psychiatrie que ne le feraient d'autres praticiens et auraient donc tendance à s'appuyer sur un projet comme les RDoC. Bien entendu, ils n'ont pas sous les yeux la matrice des RDoC, qu'ils utiliseraient plutôt comme une « boussole ». Adopter une telle conception de la psychiatrie revient à définir une espèce psychiatrique au prisme de la biologie. Il en est de même avec l'exemple du DSM. Certains praticiens comptent avant tout sur les catégories de TSA et de TDAH pour guider leur diagnostic et leur prise en charge et d'autres les utilisent de manière « opportuniste ». Concernant cette dernière utilisation, lorsqu'un doute émerge quant au diagnostic, par exemple au sein d'une équipe de soin, ce doute pourra être levé en « comptant » le nombre de critères du DSM présentés par le patient. De fait, le recours au DSM reflète l'adhésion à une certaine conception de la psychiatrie pouvant émerger dans des situations particulières au sein de certaines équipes de soin.

Le recours à telle ou telle classification, reflétant telle ou telle conception de la psychiatrie, dépendra principalement de la *fonction* (au sens de l'« objectif » ou du « but ») de la classification à un moment donné et pour un contexte précis. Par exemple, plusieurs cliniciens (ou un même clinicien dans différentes situations ou différents contextes) utiliseront un modèle fondé sur la réunion de construits psychologiques relativement larges, comme le spectre d'internalisation et d'externalisation de la HiTOP (qu'on retrouve pour une part sous le terme de structures névrotiques et psychotiques en psychanalyse),

leur permettant de cibler « grossièrement » le profil de patients pour lesquels le diagnostic n'est pas aisé, par exemple dans un contexte d'urgence. Dans d'autres contextes, d'autres cliniciens utiliseront la réunion de plusieurs catégories, comme le trouble anxieux et le trouble dépressif caractérisé, afin d'initier un traitement antidépresseur à action anxiolytique. La définition d'une espèce psychiatrique est donc utilisée selon les contraintes pratiques, dépendant des besoins cliniques, administratifs, de recherche, etc. Une question fondamentale se pose alors : en quoi la définition des espèces psychiatriques serait-elle relative aux intérêts poursuivis ? Une telle interrogation nécessite de mieux comprendre ce qu'est une espèce psychiatrique et nous y reviendrons donc par la suite dans cette thèse.

L'approche par le symptôme : présentation et intérêts

Malgré leur diversité, ces conceptions et classifications ont en commun une fragilité : elles s'appuient toutes sur des espèces diagnostiques. À ce stade, nous pouvons mieux détailler ce à quoi correspond une telle espèce diagnostique. Celle-ci renvoie à une « approche par le syndrome » (Marková et Berrios 2012). Par la suite, nous parlerons d'ailleurs plutôt d'approche « par la catégorie, dimension et spectre » – et même d'« approche par la catégorie », pour en simplifier la lecture. Cette approche par la catégorie conçoit les espèces psychiatriques comme des entités dont la validité ne peut être analysée qu'après avoir été « construites » (Cronbach et Meehl, 1955 ; Messick, 1989). Autrement dit, la catégorie est « construite » dans un premier temps et elle est validée dans un second temps. Ces temps sont ceux de l'histoire de la psychiatrie, s'étendant sur plusieurs années ou décennies. Nous verrons dans ce travail que l'approche par la catégorie ne parvient pas à remplir certaines fonctions propres à la discipline – et notamment des fonctions cliniques, qui ne peuvent que difficilement s'appuyer sur un construit préformé

en amont de la pratique clinique⁸. Ne considérer que cette approche par la catégorie, ce serait passer à côté de l'important tournant qu'a connu la recherche en psychiatrie au cours de la dernière décennie. Celle-ci s'est de plus en plus intéressée à ce qu'on peut appeler une « approche par le symptôme⁹ ». Nous verrons par la suite en quoi l'approche par le symptôme pourrait être plus pertinente pour les fonctions cliniques de la psychiatrie. En effet, une telle approche cherche à organiser les symptômes les uns par rapport aux autres, comme nous le détaillerons largement dans la seconde partie de notre thèse.

Par la suite, nous fournirons différents arguments témoignant de l'intérêt d'une telle approche par le symptôme pour le clinicien dans sa pratique. Nous avons dit que, pour une grande part, l'entretien clinique est destiné à recueillir des symptômes, mais sans nécessairement que le recours à la catégorie soit (immédiatement) nécessaire. La question de la catégorie, tout comme celle de la différenciation entre normal et pathologique, ne se pose, de manière facultative et non systématique, qu'après ces étapes essentielles de recueil et d'organisation des symptômes les uns par rapport aux autres. Ainsi, un premier argument pour adopter une approche par le symptôme tient au rôle que cette approche pourrait jouer dans la compréhension des patients rencontrés en pratique clinique. Un deuxième argument tient à un étonnement : celui de la négligence du symptôme dans la littérature contemporaine en psychiatrie et en philosophie de la psychiatrie. Cette négligence n'est-elle pas liée au fait que les cliniciens cherchent à remonter à la cause commune des symptômes recueillis plutôt que d'en chercher les relations mutuelles ? La recherche de cette cause commune permet certes aux cliniciens de pouvoir comparer des ensembles de patients dans une logique de généralisation, mais il ne s'agit plus là d'une question de « pratique

⁸ D'un point de vue méthodologique, nous avons la conviction que la validité de *construit* (que nous définirons) ne peut être pertinente que si la validité de *contenu* a été préalablement analysée – ces contenus étant des symptômes témoignant d'un préjudice, dont le rôle (mais pas spécialement la nature) devrait être nécessairement identifié. La validité de contenu a été définie comme « l'univers de problèmes qui intéresse le chercheur » (Cronbach et Meehl, 1955). La validité de contenu nous intéresse dans ce travail car elle interroge « ce que contient » une espèce psychiatrique, et répond à une question du type : « Pourquoi y-a-t-il ces symptômes-là et pas d'autres ? ». En effet, si aucune cause commune ni dysfonction n'est retrouvée à l'origine d'un « agrégat » de symptômes, pourquoi un tel construit serait utilisé et pas un autre ?

⁹ À noter que par commodité, dans cette thèse, nous regroupons les symptômes, les signes, les facteurs de risque, les pensées, émotions, comportements, expériences somatiques et éléments sociaux et contextuels sous le terme de « symptômes ».

clinique » au sens de la pratique clinique de consultation. Cette volonté de généralisation pourrait être liée à la croyance qu'une approche par la catégorie, fondée sur une cause commune, serait pertinente en pratique clinique. Une telle croyance minimiserait l'importance donnée au symptôme en tant que tel et en tant qu'élément structurant le tableau clinique *indépendamment* d'une éventuelle cause commune. Nous reviendrons largement sur ce point, car ces considérations renvoient par exemple au modèle biomédical et à l'idéal d'un modèle monocausal qui serait (« utopiquement ») utile. Un troisième argument pour adopter une approche par le symptôme tient à la difficulté conceptuelle de concevoir les espèces psychiatriques comme des espèces naturelles – et par extension comme des catégories. Il n'est pas évident que les espèces psychiatriques puissent être résumées à des entités similaires en fonction des patients, ou à des essences indépendantes du clinicien. En pratique, des décennies de recherche ont buté sur l'identification de telles essences. Les maladies paradigmatiques comme la maladie de Lyme ou la tuberculose, inscrites profondément dans l'histoire moderne de la médecine occidentale, sont bien plutôt des exceptions que des régularités. Ce troisième argument, rejetant la conception des troubles psychiatriques en termes d'espèces naturelles, nous conduit à cette prise de position épistémologique qui consiste à concevoir les troubles sur un plan sémiologique, c'est-à-dire, plus exactement, à les concevoir en termes de constellations de symptômes (ou consistant en de telles constellations). Aucune cause commune, trait commun, ni essence sous-jacente à ces réseaux ne semble utile pour le clinicien¹⁰. L'espèce psychiatrique pertinente pour le clinicien n'est pas une « variable latente ». Une variable latente (du latin, participe présent de *lateo*, « mensonge caché ») pourrait être définie comme toute variable psychiatrique qui ne peut être déduite qu'indirectement à partir d'autres variables observables ou mesurables¹¹. Une variable latente sert à regrouper plusieurs symptômes en une seule entité sous-jacente. La variable latente est utilisée pour rassembler plusieurs éléments. Par exemple, il peut s'agir d'un

¹⁰ Nous nuancerons cette position en montrant que dans certaines situations cliniques, une variable latente peut cependant être utile.

¹¹ En un sens, discuter la notion de variable latente revient pour une part à discuter le modèle médical, qui sous-tend une entité unique sous-jacent aux troubles psychiatriques. Dire cela n'est pas faux, mais nous verrons dans cette thèse qu'aborder la question au prisme de la variable latente est plus fructueux.

ensemble de traits communs, d'une cause commune, d'une catégorie diagnostique nominale (la « schizophrénie »), d'un construit statistique ou d'une dysfonction biologique. Nous soutenons que la difficulté à définir ce qu'est une telle variable latente pourrait être à l'origine d'un grand nombre de désagréments retrouvés en pratique clinique, dans la recherche psychiatrique et en philosophie de la psychiatrie. Nous rejetons la possibilité que de telles constellations de symptômes soient elles-mêmes des espèces naturelles, puisque ces ensembles de symptômes forment des réseaux directement manipulables par le clinicien. Cet intérêt pour une approche se détournant des variables latentes montre notre intérêt pour l'explicitation du contour de ce qui est pensé par le clinicien lorsqu'il recueille des symptômes : nous nous penchons en effet sur une espèce psychiatrique dans l'objectif d'en proposer une définition utile pour le clinicien. Il s'agit de spécifier les conditions d'attribution d'un trouble, c'est-à-dire spécifier à partir de quand le clinicien est amené à diagnostiquer et traiter.

Nous soutenons donc une première thèse, à portée très générale, selon laquelle la définition de ce qu'est une espèce psychiatrique est utile pour la pratique clinique. Nous argumenterons que plusieurs définitions peuvent même être nécessaires. Ce pluralisme de définitions découle d'un pluralisme de classifications, étant donné qu'une classification renvoie à une définition précise d'une espèce psychiatrique. Nous étudierons les fonctions des classifications et les fonctions des définitions qui leur sont associées. Nous poserons alors la question suivante : en quoi une définition spécifiquement clinique de ce qu'est une espèce psychiatrique peut-elle être fructueuse pour la discipline ?

Thèse principale

Notre thèse principale est donc celle-ci : la multiplicité des conceptions de la psychiatrie, instanciées au sein des classifications, serait liée à la volonté de répondre à la fois à des fonctions cliniques

(diagnostiques, pronostiques et thérapeutiques) et non-cliniques (de recherche, de santé publique, institutionnelles, politiques, financières, etc.). Plus précisément, la multiplicité des conceptions et des classifications psychiatriques serait la conséquence d'une volonté de réduction de cet écart, d'un « match » entre les fonctions cliniques et non-cliniques de la psychiatrie. La diversité de ces conceptions et classifications en psychiatrie découlerait d'une tentative visant à une concordance entre les aspects cliniques et non cliniques de la psychiatrie. Il existerait en effet une tension, au sein de chaque conception de la psychiatrie, entre la volonté d'être pertinent pour des fonctions cliniques et celle d'être pertinent pour les fonctions non-cliniques¹². Cette tension aurait conduit les différentes conceptions et classifications de la psychiatrie à adopter une approche par la catégorie. Nous soutiendrons que l'accent mis sur les catégories et dimensions dans les recommandations cliniques, les interventions pharmacologiques, le codage administratif ou la pédagogie a minimisé pour une part l'importance d'une définition utile pour le clinicien, dans sa clinique quotidienne et pour un patient donné. Nous ne souhaitons pas réduire cet écart entre fonctions cliniques et non-cliniques, mais plutôt l'accepter. Nous répondons à cette tension (ou « tentative de concordance ») en cherchant un moyen rigoureux de l'assumer, plutôt que de tenter de la contourner ou de la résoudre¹³. Nous verrons que le respect tolérant de cet écart entre fonctions est nécessaire à l'équilibre de la discipline. La possibilité d'assumer cet écart passe par l'acceptation de différentes définitions de ce qu'est une espèce psychiatrique. Nous proposons ainsi de renoncer à une unique définition d'une espèce psychiatrique ayant l'ambition d'articuler fonctions cliniques et non-cliniques. Pour cela, nous soutenons qu'au moins deux types d'espèces psychiatriques – renvoyant à l'approche par la catégorie et à l'approche par le symptôme et ayant des fonctions différentes – pourraient coexister pour garantir la stabilité de la discipline.

¹² Dans ce travail, nous faisons donc la différence entre : des conceptions (de la psychiatrie), des classifications (psychiatriques), des espèces (psychiatriques), des approches (par la catégorie/variable latente ou par le symptôme) et des fonctions (clinique et non-cliniques).

¹³ Toute classification comporte certes des fonctions cliniques et non-cliniques et, d'ailleurs, une classification ne semblerait que peu légitime si toutes ses fonctions cliniques étaient réduites à des fonctions non-cliniques. Mais nous soutiendrons ici que les fonctions cliniques

Le premier type d'espèce psychiatrique correspond à l'espèce diagnostique, qui renvoie à l'approche par la catégorie. Celle-ci sert à déterminer quelles conditions peuvent être considérées comme pathologiques, avec toutes les conséquences administratives, financières, politiques, nosologiques, pour la recherche, l'organisation sociale du système de santé et la santé publique, les politiques de soins et les recommandations de bonnes pratiques et thérapeutiques auxquelles une telle définition peut répondre (Frances 2013 ; Spitzer 1983). Elle peut être utile, concrètement, pour développer des recommandations de bonnes pratiques, de soutenir la rédaction de projets de recherche ou de remboursement des soins, d'aider au codage de la CIM ou de promouvoir la recherche en santé publique. L'analyse de la dysfonction préjudiciable (*Harmful Dysfunction Analysis – HDA*) de Jerome Wakefield fournit une définition pour une telle espèce psychiatrique ayant la capacité d'être une plaque tournante (ou « *hub* ») épistémique (Keuck 2011 ; Wakefield 1992). Un *hub* épistémique correspond à un outil scientifique (ici, une définition) qui centralise des perspectives de connaissances issues de multiples sources et disciplines (scientifiques). En servant de « concentrateur » pour répondre à différentes fonctions, le *hub* que constitue une telle définition facilite les échanges, la communication entre différentes institutions (par exemple, les institutions médicales et les communautés de patients) et différentes disciplines (par exemple, la médecine et les neurosciences). Cette définition se fonde sur l'approche par la catégorie ou dimension.

Le second type d'espèce psychiatrique mobilise une définition théorico-clinique, au sens où elle est prioritairement utile au clinicien. Le clinicien collecte des symptômes et préjudices dans sa pratique quotidienne (une tristesse de l'humeur, des hallucinations acoustico-verbales, un sentiment de jalousie, un délire mégalomane, etc.). En reprenant l'exemple de la HDA en tant qu'exemple type d'approche par la catégorie, fondée sur une espèce diagnostique, nous verrons qu'avancer une telle définition théorico-clinique semble nécessaire. Quelles preuves empiriques permettraient en effet d'affirmer que les cliniciens supposent l'existence de l'échec d'un mécanisme à exercer la fonction naturelle qui a été façonnée par l'évolution (De Vreese 2021) ? De même, un clinicien a-t-il déjà vu un patient en

consultation qui se plaignait de symptômes tel qu'un « préjudice lié à un dysfonctionnement suspecté » ? Comment le clinicien parviendra-t-il à penser une dysfonction agrégeant l'ensemble des symptômes recueillis et, *a fortiori*, un préjudice unique relié à cette dysfonction agrégeant également ces symptômes ? À travers de multiples exemples cliniques, nous verrons que la définition fondée sur la HDA n'est pas directement utile pour le clinicien. Wakefield a pourtant soutenu que sa définition, qui repose sur une catégorie avec variable latente, pourrait être utilisée dans des contextes à la fois non-cliniques (pour la recherche et la construction des nosographies) et cliniques (par exemple, pour aider le clinicien dans sa pratique)¹⁴. Nous soutiendrons dans cette thèse que le clinicien suppose plutôt l'existence de mécanismes (activés) rendant compte des symptômes (préjudiciables) qu'il recueille – définition d'un espèce psychiatrique renvoyant à une approche par le symptôme. Nous défendrons que la conception théorico-clinique du trouble psychiatrique renvoie à un ensemble d'activations de symptômes préjudiciables en interactions mutuelles. Pour la prise en charge diagnostique et thérapeutique, la réorientation, le conseil, l'accompagnement, la prévention des risques, l'élimination des diagnostics différentiels, le clinicien a besoin d'une telle définition manipulable et intuitive. Afin de fournir une telle définition aux cliniciens, nous tentons de déconstruire deux présupposés.

Le premier présupposé concerne la croyance erronée des cliniciens considérant les troubles psychiatriques comme des entités aux frontières nettes, sous-tendues par une variable latente (par exemple, un dysfonctionnement unique). Cette considération n'est que rarement explicite chez les cliniciens, mais est relativement fréquente. En effet, certains cliniciens essaient de ramener la symptomatologie à une unité (la variable latente) afin de s'approprier un diagnostic manipulable, permettant un diagnostic positif (résumer le patient à un trouble), un diagnostic différentiel (comparer

¹⁴ Même si Wakefield n'avait pas soutenu la pertinence de sa définition pour la clinique, c'est justement sa définition qui est avancée par certains cliniciens qui tentent de s'appuyer, plus ou moins explicitement, sur la notion de dysfonction et de préjudice pour comprendre leurs patients. Robert Spitzer en est un exemple. C'est ce genre de tentatives malvenues (vouées à l'échec selon nous) que nous tenterons de réfuter dans la suite de ce travail.

une diagnostic avec d'autres) et un diagnostic étiologique (réduire le trouble à une entité)¹⁵. C'est donc en partie ce réalisme naïf, la croyance en un naturalisme des espèces sous-jacentes à une variable latente, qui nous semble critiquable chez ces cliniciens. C'est la déconstruction de ce premier présupposé qui ouvre la voie à la réflexion en termes d'approche par le symptôme.

Le deuxième présupposé concerne la manière dont la philosophie de la psychiatrie a abordé la question du préjudice. Celui-ci a été principalement traité au travers des espèces diagnostiques, c'est-à-dire des catégories et au prisme des variables latentes. Nous supposons plutôt que les cliniciens perçoivent les préjudices vécus par les patients principalement à travers les symptômes recueillis en clinique. Le **Tableau 1** fournit un aperçu des différentes dénominations renvoyant à une approche par la catégorie, à différents niveaux de description.

Tableau 1 – Caractéristiques de l'approche par la catégorie. Nous rappelons que nous ne cherchons pas à fournir une définition spécifique d'une espèce psychiatrique en fonction d'une classification ou conception de la psychiatrie.

Niveaux de description	Approche par la catégorie / dimension
Entité	Variable latente
Espèce psychiatrique	Espèce diagnostique
Conceptions	DSM, HiTOP ou projet RDoC (« modèle médical »)

¹⁵ Notons que dans le débat sur les espèces psychiatriques, il semble nécessaire de bien différencier la question du normal et du pathologique (par exemple : « Les symptômes appartenant à la catégorie de la dépression constituent-ils ensemble un trouble psychiatrique ? ») de la question de dimensionnalité (par exemple : « Quel degré de tristesse est présent chez ce patient avec un tableau clinique faisant référence à la catégorie de la dépression ? ») et du diagnostic différentiel (par exemple : « En quoi ces symptômes appartenant à la catégorie de la dépression se distinguent-ils de ceux de l'anxiété généralisée ? »). Dans cette thèse, fondée sur la notion de symptômes préjudiciables, nous nous intéressons à la première de ces questions, à savoir la différenciation entre normal et pathologique.

Exemple	Le trouble de la communication sociale pragmatique est défini selon une dysfonction dans la communication sociale entraînant un préjudice pour la qualité des interactions de l'individu (exemple de l'analyse de la dysfonction préjudiciable)
---------	---

Tolérer l'existence d'au moins deux types d'espèces psychiatriques implique d'accepter nécessairement une unité « bipartite » pour la discipline. Issu du vocabulaire des mathématiques (et avant cela, de la théologie), un ensemble bipartite est un ensemble d'éléments qui est séparé en deux – il s'agit d'un ensemble unique contenant deux types d'éléments. L'intérêt d'employer cette notion de « bipartisme » est de rappeler qu'une unité pour la discipline est toujours possible même si elle comprend des définitions distinctes – c'est en tout cas un des principaux enjeux de cette thèse. En termes philosophiques, il s'agirait donc d'un pluralisme définitionnel, bipartite, permettant à la psychiatrie de présenter deux types de définitions tout en conservant son unité. Nous ne discuterons cependant pas de la notion d'unité de la discipline psychiatrique, qui relève de facteurs institutionnels, sociaux, politiques, économiques et cliniques. Ainsi, parler d'une unité bipartite permet de proposer une position tempérée entre un dualisme opposant radicalement deux définitions, tout en veillant à distinguer deux ensembles distincts au sein de la psychiatrie déjà constituée en tant que discipline. Enfin, une telle unité bipartite ne nécessite pas d'opposer une vision scientifique de la psychiatrie, fondée sur la médecine factuelle, à une psychiatrie axée sur la clinique, fondée sur la subjectivité du clinicien. Notamment, nous ne cherchons pas à argumenter l'existence d'une « exclusivité » de la clinique par rapport à des fonctions non-cliniques (politiques ou administratives) de la psychiatrie (Demazeux, 2019). Les fonctions cliniques sont développées au sein de la pratique clinique, mais sont largement teintées de dimensions et de facteurs non sémiologiques ; le clinicien ne pratique pas que de la clinique « pure » et remplit d'ailleurs de nombreuses autres fonctions en parallèle. Nous choisissons de traiter les fonctions cliniques séparément des fonctions non-cliniques uniquement pour réduire l'ampleur de ce travail et parce que les fonctions cliniques nous semblent importantes en psychiatrie pour la discipline médicale qu'est la psychiatrie. Un

travail du même type pourrait être effectué pour étudier, par exemple, la différence de définition d'une espèce psychiatrique entre des fonctions politiques et non-politiques.

Enfin, un des risques apparent de cette thèse serait de faire face à cet argument : toutes les classifications ont-elles eu pour souci de répondre à une fonction clinique ? Toutes ont-elles eu la nécessité de s'adresser à la pratique clinique ? Et si c'était le cas, quel serait le lien entre cette fonction clinique et les fonctions non-cliniques ? C'est tout l'enjeu des trois premiers chapitres de cette thèse de répondre à ces questions.

De plus, deux nuances sur la séparation entre fonction clinique et fonction non-clinique pourraient être données à ce stade. En effet, une première question qui pourrait se poser serait celle-ci : peut-on seulement dissocier des fonctions cliniques d'autres fonctions non-cliniques ? L'acte diagnostique est éminemment social et institutionnel (sinon politique) en même temps que clinique. Isoler complètement la clinique d'autres enjeux sociaux, administratifs ou épistémiques et théoriques serait un contresens profond de cette thèse. Il est très important de concevoir que par « fonction clinique », nous entendons un ensemble d'objectifs requis par la pratique clinique et qui s'imposent au clinicien pour qu'il puisse exercer *face à un patient* ; il ne s'agit donc pas que d'un ensemble sémiologique « pur », détaché de toute contrainte externe à la relation médecin-patient. La notion de « pratique clinique » devrait être distinguée de celle de « sémiologie ». Une fonction clinique pourrait être définie comme un recueil et une structuration des symptômes du patient afin de lui proposer des réponses pratiques (sur les plans diagnostiques, pronostiques, préventifs et thérapeutiques). Ce recueil et cette structuration sont soutenus par de multiples facteurs liés au bagage théorique, sociologique et aux habitudes du clinicien. Nous insistons sur le fait que les « fonctions cliniques », si elles sont certes principalement axées autour du diagnostic, du pronostic et de la prise en charge (qui constituent trois objectifs de soin) ne peuvent *jamais* être détachées des impératifs culturels, sociaux, économiques ou historiques dans lesquels est pris le clinicien : les fonctions cliniques sont intimement reliées aux valeurs d'une société exerçant une pression pour médicaliser, par exemple, les refus scolaire anxieux ou les enfants à « haut potentiel intellectuel ». Dans la pratique clinique, les cliniciens s'appuieront sur la sémiologie mais considérer

également les conséquences pratiques, les stratégies médicales, les orientations de prise en charge – et plus généralement l’implémentation des politiques de santé et les orientations de soin des institutions – de telle ou telle conception de la psychiatrie, véhiculant telle ou telle définition. De même, dans leur pratique, les cliniciens considèrent des facteurs issus de la recherche, qui transforment donc leur pratique en influençant leurs raisonnements, de la même manière que peuvent le faire, par exemple, divers déterminants sociaux. Répondre à une fonction clinique nécessite d’intégrer des dimensions qui relèvent de bien d’autres dimensions (sociales, institutionnelles, etc.). Mais c’est pourtant pour répondre aux besoins de la pratique clinique – à savoir des besoins diagnostiques, pronostiques et thérapeutiques – que nous parlons de fonction clinique. Nous assumons à ce stade ne pas rigoureusement définir ce qu’est précisément une « fonction clinique » puisqu’elle s’inscrit dans de multiples ancrages qui dépassent largement la sémiologie.

Une deuxième nuance à apporter sur la différence entre fonction clinique et fonction non-clinique est celle-ci : le clinicien ne s’appuie sur les fonctions cliniques que pour une part restreinte de son activité. Il ne faut pas confondre fonction clinique et fonction du clinicien. Certes, face au patient, le clinicien cherche à recueillir des signes et symptômes, les labelliser et les organiser. Comme nous venons de le dire, ces actes nécessitent de convoquer de multiples dimensions qui dépassent largement la sémiologie. Mais en plus de cela, en dehors de l’entretien face à un patient, le clinicien s’occupe de fonctions administratives et institutionnelles (qui concernent plus ou moins directement son patient). Il répond aussi à des fonctions de santé publique et même politiques, organisationnelles et médiatiques (qui concernent l’ensemble de ses patients). Les fonctions non-cliniques de la psychiatrie concernent aussi bien le clinicien que les autres types d’intervenants qui appartiennent de près ou de loin à la discipline psychiatrique (travailleurs sociaux, politiques, administrateurs, décideurs, gestionnaires, patients eux-mêmes, etc.). La fonction clinique n’est donc jamais indépendante d’autres fonctions. Ce que nous questionnons, dans ce travail, c’est plutôt l’hégémonie de l’approche par la catégorie dans le cadre de la pratique clinique.

Plan

Cette thèse est composée de deux parties. La première permet de décrire trois différentes conceptions de la psychiatrie en étudiant leurs « cristallisations » au travers de trois classifications psychiatriques. Afin de soutenir notre thèse principale (selon laquelle la multiplicité des conceptions de la psychiatrie, instanciées au sein des classifications, serait liée à la volonté de répondre à la fois à des fonctions cliniques et non-cliniques, thèse ayant conduit différentes conceptions de la psychiatrie à adopter une approche par la catégorie), nous utilisons la nosologie comme un « outil » de discussion philosophique. Celle-ci nous permet de discuter plus précisément de l'approche par la catégorie. Le recours à la nosologie a pour fonction de présenter les différentes conceptions de la psychiatrie telles qu'elles sont déployées dans la pratique, de souligner leurs limites et de montrer que toutes les classifications s'appuient sur une volonté de satisfaire au moins des intérêts cliniques – en se fondant sur la notion de variable latente. Que la conception de la psychiatrie soit instanciée au travers des catégories du DSM, des fonctions psychologiques de la matrice des RDoC ou d'un ensemble de construits statistiques comme la HiTOP, toutes s'appuient sur la notion de variable latente. Trois défis pourraient être soulevés : 1) celui d'expliquer pourquoi de multiples fonctions se sont révélées nécessaires pour la psychiatrie, ce que nous ferons en travaillant autour de l'histoire de la constitution du DSM ; 2) celui d'expliquer comment la psychiatrie a cherché à « réconcilier » la pratique clinique et les enjeux scientifiques, au travers de l'analyse de la portée clinique et environnementale des RDoC ; 3) celui d'expliquer la volonté de se détacher des catégories en proposant une approche par les dimensions, en étudiant la structure de la HiTOP. Dans les trois premiers chapitres, nous illustrons ainsi les multiples fonctions des classifications psychiatriques via le DSM, la délicate relation entre les enjeux cliniques de la nosologie et leurs fondements scientifiques via les RDoC et les croisements entre la clinique et la recherche scientifique

(et statistique) via la HiTOP. Cette première partie de la thèse est particulièrement importante pour comprendre les enjeux d'une remise en question des catégories. Sans une connaissance structurale, historique et pratique des classifications et donc des catégories et éléments qui les composent, il semblerait ambitieux de proposer une nouvelle conception de la psychiatrie fondée sur les symptômes. Dans la seconde partie de notre thèse, nous utilisons la définition de ce qu'est une espèce psychiatrique. Nous soutiendrons une approche par le symptôme utile au clinicien. Afin de répondre à la fois à cette diversité et à cette mise en tension, nous explorons comment le symptôme peut aider à redéfinir les espèces psychiatriques. La **Figure 1** synthétise les différents termes utilisés dans les chapitres de la thèse.

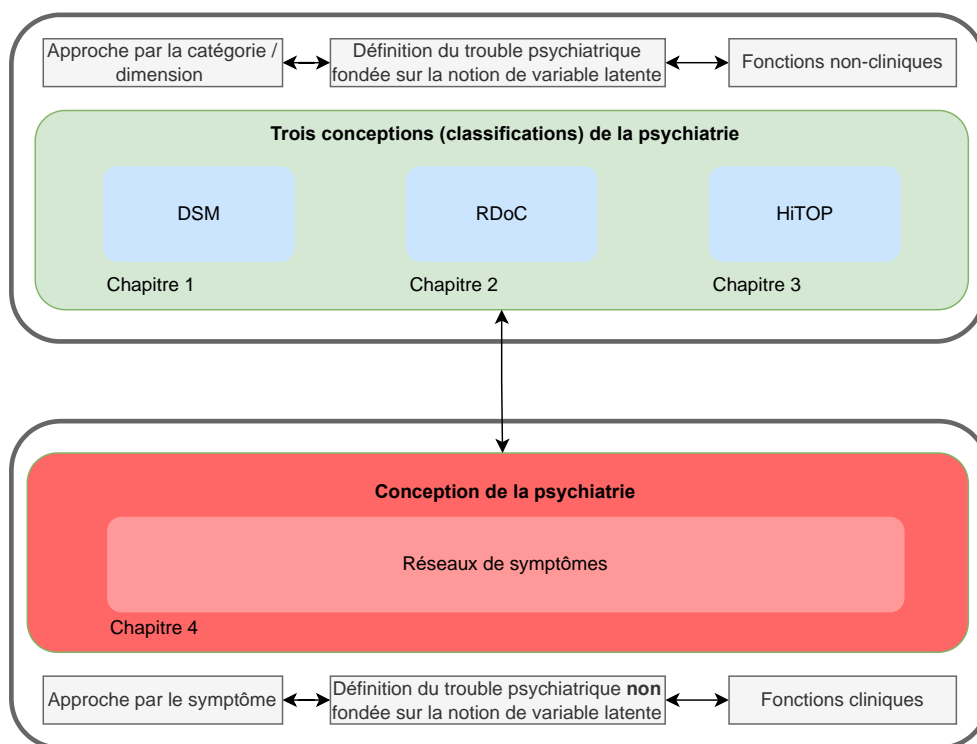


Figure 1 – Synthèse introductive de l'organisation des chapitres de la thèse. DSM : *Diagnostic and Statistical Manual of Mental Disorders* ; RDoC : *Research Domain Criteria* ; HiTOP : *Hierarchical Taxonomy Of Psychopathology*. Cette figure fournit une visualisation synthétique de ce travail et ne peut rendre compte de certains éléments de nuance, comme le fait que le DSM, les RDoC et la HiTOP remplissent également des fonctions cliniques.

PARTIE I

LES MULTIPLES FONCTIONS DU MANUEL DIAGNOSTIQUE ET STATISTIQUE DES TROUBLES MENTAUX : UN REFLET DE SON HISTOIRE

Dans ce premier chapitre, nous faisons l'hypothèse que l'étude du DSM permet d'avancer sur la compréhension de la pratique clinique de la psychiatrie contemporaine. La thèse de ce chapitre est de démontrer que la multiplicité des fonctions cliniques et non-cliniques du DSM est offerte par son approche par la catégorie et que cette approche est elle-même fondée sur un ensemble de notions (par exemple, de fiabilité ou d'usage que l'histoire de la classification met en avant).

Nous étudions l'importance des multiples utilisations du DSM dans l'histoire de la psychiatrie. Nous discutons brièvement de l'ère pré-DSM-III (avant 1980) et de l'importance des critères de Feighner et des critères diagnostiques pour la recherche. Nous soulignons l'importance des multiples racines épistémologiques de cette classification, son contexte d'instauration dans le paysage de la psychiatrie et les étapes de son développement jusqu'au DSM-5, en 2013. Cependant, notre objectif n'est pas de fournir ici une énième déclinaison du développement chronologique des versions du DSM. Plutôt, nous tissons un lien fort entre ce développement et les implications de la nosologie avec les fonctions cliniques et non-cliniques que requiert la psychiatrie.

Nous retraçons comment le DSM s'est construit sur la base du modèle médical (ce que son histoire permet de justement montrer) pour répondre à des enjeux à la fois cliniques et non-cliniques. Ce modèle médical correspond à une approche par la catégorie (expression utilisée en contraste avec l'approche par le symptôme que nous développerons par la suite), appuyée elle-même sur les notions de fiabilité ou d'opérationnalisation. Un des principaux résultats de ce chapitre est de montrer que cette approche par la catégorie, utile à de nombreuses fonctions non-cliniques, est fortement mise en tension avec la volonté de donner une pertinence clinique au DSM.

Introduction

La question de l'évolution des versions du DSM et de son ancrage contextuel paraît particulièrement importante pour comprendre la place centrale que cette classification occupe dans le paysage de la psychiatrie contemporaine. Le DSM est régulièrement présenté en recherche (et en clinique) comme un outil de travail solidement ancré dans les pratiques. L'exploration historique et philosophique de son contexte d'apparition et de développement semble essentiel pour la psychiatrie. En effet, l'analyse de la genèse de la centralité du DSM en psychiatrie éclaire les discussions sur les hypothèses philosophiques qui sous-tendent la discipline. Si le DSM n'est pas qu'un objet actuel, mais un projet dynamique fondé en partie sur l'empirisme logique de Hempel (ce que nous critiquerons en prenant en compte ses autres fondements épistémologiques), sa place dans le paysage psychiatrique contemporain pourrait être relativisée. En d'autres termes, la philosophie a beaucoup à apporter à cette possibilité de « décaler » les intérêts et les passions qui entourent le DSM. Nous verrons que ces passions sont en partie liées à la volonté des promoteurs du DSM que ce dernier soit ancré à la fois dans la pratique de recherche et dans l'épidémiologie, en même temps qu'il se veut (indirectement et presque « passivement ») constituer un manuel clinique. Nous montrerons que cette volonté d'ancrage dans différents champs n'est pas indépendante des propositions de nosographies alternatives au DSM. « L'objet impur » que constitue le DSM, selon les termes de Sadler (qui soulignait par-là ses racines diverses dans l'histoire de la discipline et le développement des statistiques en psychopathologie de la première moitié du XX^e siècle (Sadler, 2005)), rassemble au sein d'un même ouvrage des critères tendant vers l'opérationnalisation, une volonté d'athéorisme et une conceptualisation catégorielle de la psychiatrie.

Dans ce chapitre, nous soutiendrons la thèse d'une multiplicité de fonctions du DSM, qui est offert par son approche par la catégorie. Cette approche est elle-même fondée sur au moins trois traditions épistémologiques à l'origine des différentes versions du DSM : i) ces dernières sont le reflet d'une

histoire ancienne de la médecine, fondée sur la volonté de classer les maladies ; ii) elles s'appuient sur des critères diagnostiques ; iii) et elles témoignent d'une tradition psychiatrique et philosophique issue de l'opérationnalisme. Nous détaillons brièvement ces trois points ci-dessous avant de les aborder plus précisément à un niveau historique au fil de ce chapitre. Concernant le DSM en tant que reflet d'une histoire ancienne de la médecine, nous pourrions ancrer cette classification dans l'historiographie des manuels de diagnostics médicaux (Decker, 2013). Par exemple, la volonté de créer des classifications médicales pourrait être retrouvée dans un traité comme le *Nosologia Methodica* de François Boissier de Sauvages de Lacroix. En 1763, ce médecin et botaniste de Montpellier, inspiré par la classification des plantes, présentait dix classes de maladies dont la huitième était la folie et où chaque classe était divisée en genres et en espèces (Boissier de Sauvages de Lacroix, 1763). Concernant le DSM en tant que manuel structuré autour de critères diagnostiques, la volonté de proposer des critères rigoureux et systématiques pourrait renvoyer aux programmes de recherche débutés avant les années 1950, notamment par des psychiatres formés en méthodologie, en statistique et à la psychométrie (Goodwin et al., 1969; Purtell et al., 1951; Robins & Menseh, 1954). Ces derniers cherchaient à encourager la recherche et la méthodologie en psychiatrie, notamment pour réfuter les affirmations d'antipsychiatres pour qui le trouble mental serait une « fiction théorique » (Szasz, 1960). En plus de cela, au début des années 1950, un groupe de psychiatres cliniciens se rassemblaient autour de l'idée du diagnostic, prônant une approche de la psychiatrie tournée vers le recueil systématique d'éléments sémiologiques agrégés (ce point est important pour comprendre la seconde partie de la thèse, fondée sur l'idée d'une approche par le symptôme). Le DSM est enfin le reflet d'une tradition psychiatrique et philosophique. Les critères de Feighner et les critères diagnostiques pour la recherche, que nous développerons plus bas, s'appuient justement sur la notion philosophique d'opérationnalisation. Celle-ci aide à considérer de quelle manière une classification, en se montrant « manipulable », peut être robuste, c'est-à-dire (idéalement) à la fois fiable, valide, claire, parcimonieuse, reproductible, intégrative, cliniquement pertinente, acceptable, sensible, spécifique et utile à différentes fins (cette liste de termes est employée indistinctement et se recoupe dans

la littérature psychiatrique et de philosophie de la psychiatrie, sans qu'un terme ne recouvre tous les autres). Dans ce chapitre, en nous référant à une abondante littérature sur l'histoire et l'élaboration des différentes versions du DSM, nous aborderons les fonctions de cette classification selon un plan chronologique – afin de simplifier la présentation des arguments de la thèse du chapitre (selon laquelle la multiplicité des fonctions du DSM est rendue possible par la volonté de créer une approche par la catégorie). Nous traiterons dans une première partie de l'ère pré-DSM-III (avant 1980). Dans une deuxième partie, nous soulignerons l'importance des critères de Feighner et des critères diagnostiques pour la recherche, pour aborder dans une troisième partie le processus et les conditions de développement du DSM-III. Nous traiterons dans une quatrième partie les processus de révision et d'élaboration des différentes versions du DSM.

I. En amont du DSM

À la fin du XIX^e siècle, en Europe, les approches diagnostiques en psychiatrie ont commencé à se diviser en au moins deux écoles distinctes. D'un côté, Freud et la psychanalyse soulignaient l'importance de « l'étiopathogénie psychologique » des troubles mentaux et les inscrivaient dans un continuum avec le fonctionnement normal (Sandford, 2017). D'un autre côté, un groupe de psychiatres, dont faisait partie Emil Kraepelin, cherchait à appliquer un modèle médical qui examinait les symptômes, l'évolution et le pronostic des troubles (Engstrom & Kendler, 2015). Il faudra souligner l'importance de l'approche kraepelinienne dans l'incitation des chercheurs à développer un tel programme de recherche qui utilise des techniques statistiques pour cerner les troubles psychopathologiques – programme qui conduira à la constitution du DSM-III. Dans cette partie, en partant de la pensée kraepelinienne, nous retracerons l'histoire des principaux concepts et valeurs qui ont été véhiculés au sein des différentes versions du DSM.

Avant Kraepelin, en médecine, les symptômes permettaient d'individualiser les troubles et maladies. Cependant, en l'absence de principe communément partagé sur la manière de diviser les diverses symptômes perçus par les médecins, la nécessité d'une logique regroupant les symptômes s'est fait ressentir. Cette logique de regroupement a montré l'importance de discuter de syndromes (c'est-à-dire justement des regroupements de symptômes). Sans entrer dans les différentes méthodes de classifications que nous détaillons ailleurs¹⁶, nous pourrions en observer les résultats, tels qu'ils sont apparus dans différents systèmes à des moments précis de l'histoire de la psychiatrie (Ellenberger, 1981; Heckers & Kendler, 2020; Hoff, 2015; Kendler & Engstrom, 2018; Leonhard, 1999). À un extrême, ces méthodes de classification ont conduit à considérer une multitude de troubles distincts dès lors que des symptômes différaient entre eux – conduisant à une prolifération de troubles (pouvant atteindre des centaines de catégories)¹⁷. À l'opposé, la diversité et l'hétérogénéité des tableaux cliniques a pu être considérée comme un ensemble de variantes de quelques troubles seulement. Ainsi, le premier recensement américain sur les troubles mentaux réalisé en 1840 reflétait cette dernière approche et ne contenait qu'une catégorie de trouble psychiatrique (la « folie »), établie sur des bases socio-historiques (Kendler & Parnas, 2012). Nous aborderons plus en détail ce débat en détaillant la HiTOP, qui intègre notamment dans sa hiérarchie un « facteur p », c'est-à-dire un facteur statistique renvoyant à l'ensemble des classes relevant de la psychopathologie. Classer nécessite donc d'identifier des critères précis permettant de séparer ou rassembler les troubles les uns par rapport aux autres, sur la base des symptômes.

Emil Kraepelin (1856 – 1926) a joué un rôle important dans cette tentative d'organisation sémiologique. Il a commencé sa carrière en tant que médecin dans un asile munichois et a maintenu son intérêt (presque exclusif) pour les troubles psychotiques en tant que professeur à Heidelberg et en tant que directeur de la clinique psychiatrique à Munich. Psychiatre allemand, contemporain de Freud, il a utilisé les

¹⁶ Nous étudierons largement par la suite comment ce « regroupement » a pu être considéré tantôt comme le fruit d'une validité de construit (Cronbach & Meehl, 1955), tantôt comme le noyau central prototypique que des auteurs comme Sadegh-Zadeh ont pu décrire dans le cadre de la médecine (Sadegh-Zadeh, 2008).

¹⁷ Comme nous le verrons dans une annexe sur la psychiatrie computationnelle, cette méthode de séparation des troubles (*split*) se rapproche de ce qui est fait avec les modèles computationnels qui se « sur-ajustent » à toute nouvelle donnée clinique.

symptômes et l'évolution des troubles pour proposer des catégories qui représentaient des conditions pathologiques sous-jacentes distinctes, qu'il espérait pouvoir éventuellement confirmer par l'identification de lésions anatomiques – bien que restant à distance des ambitions de la neuropsychiatrie de Meynert et Wernicke (Blashfield, 1993; Blashfield et al., 1992). La conception de Kraepelin, fondée principalement sur les symptômes et sur la possibilité d'en inférer des états pathologiques distincts, suivait une tradition médicale datant du XVIII^e siècle et portée notamment par le médecin anglais Thomas Sydenham, elle-même développée par un pathologiste allemand du XIX^e siècle, Rudolph Virchow (Demazeux, 2013a; Ghaemi, 2004). Kraepelin utilisait notamment la notion d'inférence pronostique pour faire la distinction entre la folie maniaco-dépressive (le trouble bipolaire actuellement), qui survenait par événements ponctuels et la *dementia praecox* (trouble du spectre de la schizophrénie actuellement) (Shepherd, 1987), qui avaient tendance à entraîner tous deux une dégradation de l'état du patient. Les contributions de Kraepelin au modèle médical de la psychiatrie que nous connaissons actuellement (en particulier à travers sa volonté de catégorisation fondée sur les symptômes) ont pu être considérées comme des préalables à la constitution du DSM-III¹⁸. Cette histoire complexe de la psychiatrie pré-DSM a été retracée par de multiples commentateurs et sous différents angles, tantôt en insistant sur l'importance des aliénistes de la fin du XVIII^e siècle (Goldstein, 1987), tantôt sur l'importance des statistiques et des catégories (telles que la manie, la mélancolie ou la paranoïa) (Postel et Quénel, 2012) et tantôt sur la volonté de construire une classification universelle destinée à la psychiatrie, à l'image de la CIM (Kendler & Parnas, 2017). Afin de ne pas avoir à retracer l'intégralité de ces développements et à prendre position vis-à-vis de ces sources historiques, nous nous limiterons aux conditions de développement des critères de Feighner et des critères diagnostiques pour la recherche.

¹⁸ Robert Spitzer, que nous présenterons par la suite, niera cependant être un « néo-kraepelinien », au motif qu'il « ne suppose pas qu'il doit exister des pathologies catégorielles distinctes sous-jacentes à différents syndromes », ni que « les troubles mentaux sont en grande partie dus à des troubles physiques contenus dans le cerveau » – correspondant à deux principes de base de l'approche dite kraepelinienne (Heckers & Kendler, 2020; Spitzer et al., 1975).

C'est en présentant ceux-ci que nous comprendrons mieux l'importance de la nosologie pour la psychiatrie.

II. Les critères de Feighner et les critères diagnostiques pour la recherche

La nomenclature psychiatrique aux États-Unis pendant la première moitié du XX^e siècle ne reflétait pas un intérêt intense pour la classification psychiatrique¹⁹. Ce n'est que par la suite qu'un besoin administratif croissant d'obtenir des statistiques sur les troubles (psychiques) aurait accéléré la volonté d'élaboration des manuels diagnostiques (Grob, 1985).

1. La CIM et le DSM

Nous soutenons dans ce chapitre que les multiples fonctions du DSM ont été influencées par le contexte d'instauration et de déploiement de la classification. Afin de bien comprendre le contexte dans lequel ont été publiées les différentes versions du DSM, il est nécessaire d'étudier l'importance de classifications internationales comme la Classification Internationale des Maladies (CIM). On relève en effet des tentatives de développement communes entre la CIM-6 et le DSM-I, la CIM-8 et le DSM-II²⁰,

¹⁹ Anne Fagot-Largeault a largement développé cette différence entre nomenclature (comme « ensemble de termes en usage ») et classification.

²⁰ Dans un interview original publié en 2022, Robert Spitzer affirme explicitement que le DSM-II avait pris exemple sur la CIM-8 « à la moindre virgule et parenthèse près », tout en étant l'œuvre d'un (presque unique) auteur, Sir Aubrey Lewis, sur lequel nous reviendrons par la suite (<https://www.madinamerica.com/2022/02/robert-spitzer-dsm/>).

la CIM-9 et le DSM-III et la CIM-10 et le DSM-IV, avec une harmonie et une cohérence croissantes du fait de tentatives de collaboration internationale (Johnstone et al., 2010). Dans cette section, nous tenterons d'en explorer les développements croisés. La CIM a été proposée initialement en tant que système diagnostique « universel ». Cette tentative d'universalisme nosologique a symboliquement débuté lors d'un congrès international d'épidémiologie et statistique, en 1891, avec un accord visant à établir une liste des causes de décès dans le cadre d'une utilisation internationale (Fagot-Largeault, 1991). Au cours des années 1950, la nécessité d'évaluer l'état mental d'un grand nombre de militaires, à la suite de la Seconde Guerre mondiale, a conduit à rechercher une meilleure communication entre les psychiatres, les professionnels de la santé mentale et plus généralement les autres membres du monde médico-social (Wilson, 1993). En 1949, l'Organisation mondiale de la santé publiait alors la sixième révision de la CIM, qui comportait pour la première fois une section sur les troubles mentaux. À la suite de la publication de la CIM-6, un schéma de classification psychiatrique a été proposé par l'*American Psychiatric Association* (APA), notamment pour pouvoir évaluer les militaires américains sur le plan psychique. Le Bureau du chirurgien général, nommé « Medical 203 », publiait ainsi le DSM-I en 1952 (Demazeux, 2013a). Le DSM-I était censé mieux refléter la nature de l'évolution de la population de patients de la profession psychiatrique (Decker, 2013; Horwitz, 2014). À cette époque, l'idée que les troubles mentaux résultaient de facteurs psychologiques, sociaux et biologiques était prédominante. Cependant, jusqu'au moins la CIM-9 parue en 1977, l'approche psychodynamique restait prédominante dans la terminologie et la conception des troubles (Werry, 1985). Peu de temps après la publication du DSM-I, l'OMS a chargé Erwin Stengel de passer en revue les principales classifications psychiatriques disponibles et de faire des suggestions pour améliorer la communication internationale. Stengel a illustré « l'état chaotique des classifications actuellement utilisées » en rassemblant vingt-huit classifications officielles, semi-officielles ou nationales, voire régionales (par exemple, il existait une classification spécifique pour l'État de New-York). Selon lui, aucune de ces classifications n'était accompagnée d'un glossaire suffisamment bien construit, c'est-à-dire une définition convenable des termes de la psychiatrie

(Stengel, 1959). Il en soulignait cependant l'importance, en précisant comment de telles définitions pouvaient conditionner la compréhension de la classification et donc du soin clinique (Henderson, 1950). Et en effet, la CIM-7, publiée en 1955, était simplement constituée d'une liste de noms et de numéros de code, sans indication du sens préférentiel des termes utilisés (Stengel, 1959). La CIM-8, publiée par l'OMS en 1968, proposait pour la première fois une classification internationale accompagnée d'un glossaire censé être opérationnel (nous reviendrons dans la section III sur cette notion), fourni séparément du manuel (Kendler, 2018). C'est cette huitième révision de la CIM, accompagnée de son glossaire, qui a incité l'APA à produire le DSM-II. Celui-ci a été publié en 1968 et se basait principalement sur la CIM-8 parue la même année. Comme souligné ci-dessus pour la CIM, le DSM-I et le DSM-II reflétaient l'approche psychodynamique alors prédominante aux États-Unis, sans réellement décrire de manière standardisée les symptômes des troubles psychiatriques (Bayer & Spitzer, 1985). Malgré l'existence d'un glossaire, avant les années 1970, la classification diagnostique psychiatrique était « fondée sur le meilleur jugement clinique et l'expérience d'un comité et de ses consultants » (Feighner et al., 1972). Les DSM-I et DSM-II étaient constitués de descriptions brèves et purement narratives des catégories, sans critériologie ni opérationnalisation. Ce sont (entre autres) ces caractéristiques qui ont limité leur utilisation aux États-Unis ou à l'international. Ce n'est qu'avec les versions suivantes de la CIM et du DSM que la notion d'opérationnalisation s'est un peu plus ancrée dans la nosologie. Alors que Tsung-yi, dans les années 1960, était responsable de la psychiatrie épidémiologique et sociale de l'Unité de la Santé Mentale de l'OMS, il fut chargé de constituer un programme innovant en vue de publier la CIM-9, publication qui sera effective en 1975 et publiée sous le titre évocateur de *La place de l'épidémiologie en psychiatrie* (Standley & Lin Tsong-Yi, 1963)²¹. Ce programme ne prévoyait cependant pas d'abandonner l'approche psychodynamique qui était à l'œuvre dans les versions précédentes de la CIM (Werry, 1985). De fait, malgré ces propositions apparemment

²¹ Ce programme prévoyait quatre composantes : i) développer une classification internationale des troubles mentaux, ii) établir des procédures standardisées, iii) permettre la formation d'un personnel expérimenté, iv) et mettre en place un réseau de centres de recherche et de formation.

novatrices, la CIM-9 fut publiée sans changement majeur dans sa structure, selon une approche conservatrice possédant une forte tendance psychodynamique (Kendler et al., 2010).

En parallèle de cette publication, en 1974, l'APA nommait quatorze groupes de travail²² sur la nomenclature et les statistiques pour commencer les travaux qui devraient conduire au développement du DSM-III – en cherchant à se conformer à la CIM-9 (Akram & Giordano, 2017; Spitzer et al., 1975).

En effet, au moment où ce nouveau groupe de travail de l'APA était constitué, la section de la CIM-9 portant sur les troubles mentaux était presque terminée. Or, de nombreux domaines spécifiques de la CIM-9 apparurent comme insuffisamment détaillés pour être utilisés en clinique et en recherche. De plus, contrairement à ce qui était attendu pour le DSM-III en cours de constitution, la CIM-9 ne recherchait pas de cohérence interdisciplinaire, c'est-à-dire de cohérence entre les critères psychiatriques et ceux d'autres spécialités médicales (comme la neurologie). Les quatorze groupes de travail de l'APA ont été chargés de préparer une classification reflétant mieux l'état des connaissances concernant les troubles mentaux, tout en veillant à conserver la compatibilité avec la CIM-9.

2. Les raisons de l'acceptation d'un nouveau système de classification pour la psychiatrie

Le système proposé par le DSM-III (que nous décrirons – et fondé notamment sur les critères de Feighner et les critères diagnostiques pour la recherche) répondait à plusieurs difficultés majeures auxquelles les cliniciens (comme les chercheurs) étaient confrontés à l'époque. Ces réponses constituent l'une des forces de la multifonctionnalité du DSM. Dans l'Amérique du Nord des années 1970, l'influence psychanalytique s'était quelque peu estompée. De nombreuses écoles théoriques et des cliniciens aux

²² Chacun était considéré comme un « comité consultatif », composé de professionnels ayant une expertise particulière dans chaque domaine (Demazeux, 2013a).

formations très diverses partageaient une même population clinique avec des hypothèses de causalité, des mécanismes descriptifs ou des traitements bien différents. En se plaçant dans une perspective « athéorique », le DSM-III promettait d'accueillir ces multiples perspectives (Spitzer, 1983). En se positionnant comme ensemble de *références descriptives*, plutôt qu'en renvoyant à des étiologies propres à un domaine scientifique (rendant compte de cette volonté d'athéorisme), les critères ont pu rapidement gagner en fiabilité²³ et, par extension, en popularité.

Avant la constitution du DSM-III, les diagnostics psychiatriques étaient sujets à des critiques provenant de nombreuses et diverses sources (et pas uniquement de celles issues des milieux psychanalytiques). D'un côté, les comportementalistes affirmaient que les comportements (y compris la psychopathologie) résultaient de processus d'apprentissage normaux et qu'il n'existait donc pas de trouble mental au sens médical du terme (Eysenck et al., 1983). D'un autre côté, le mouvement antipsychiatrique, inspiré par des auteurs comme le psychiatre Thomas Szasz (Szasz, 1960) ou le sociologue Thomas Scheff (Scheff, 1971) affirmaient que le diagnostic psychiatrique consistait à utiliser la terminologie médicale pour développer un contrôle social sans qu'il n'y ait réellement de trouble médical à proprement parler²⁴. De plus, aux alentours des années 1980, « l'érosion de la crédibilité » de la psychiatrie était fortement rapportée à l'impossibilité de réaliser des tests empiriques susceptibles d'assurer son statut scientifique (Bayer & Spitzer, 1985) – statut nécessaire pour médicaliser la discipline, mais utile aussi en termes de remboursement et de crédibilité politique et institutionnelle. L'expérience de 1972 de Rosenhan, qui a (à l'époque) servi à démontrer l'absence de fiabilité des diagnostics psychiatriques, a contribué à cette érosion. Rosenhan écrivait par exemple : « Fondée en partie sur des considérations théoriques et anthropologiques, mais aussi philosophiques, juridiques et thérapeutiques, nous sommes de plus en plus convaincu que la catégorisation psychologique de la maladie mentale est au mieux inutile et au pire nuisible ou trompeuse » (Rosenhan, 1973). Aujourd'hui, malgré les intenses critiques et la démonstration

²³ Nous reviendrons largement sur cette notion, qui correspond brièvement au fait que les cliniciens obtiennent un même résultat entre eux et / ou dans le temps.

²⁴ Bien que plus subtilement, Szasz ne niait pas qu'il puisse y avoir une « essence » des troubles mentaux.

des biais portant sur cette recherche (notamment développées par Cahalan dans *The Great Pretender: The Undercover Mission That Changed Our Understanding of Madness* (Cahalan, 2019)²⁵), il ne peut être nié que l'impact d'une meilleure fiabilité des troubles psychiatriques a été majeur.

Mais tous ces facteurs n'ont pas conduit à un effondrement de la psychiatrie sur elle-même : ils ont sans doute contribué à *faire le vide*, pour laisser de la place à la classification naissante que représentait le DSM-III. La question qui se pose alors est celle de savoir si le DSM-III a pu contribuer à unifier et ordonner les éléments d'un contexte psychiatrique qui apparaissait fragmenté. Cette organisation autour de la classification, conçue en tant que point de passage obligé et « manipulable » pour un ensemble d'acteurs et de communautés partageant une conception relativement similaire, constitue le cœur des questionnements de ce chapitre. Nous reviendrons sur cet intérêt par la suite, en développant notamment le contexte d'apparition des critères de Feighner et des critères diagnostiques pour la recherche. Ce n'est qu'en abordant de tels éléments historiques que nous pourrions comprendre en quoi le DSM est devenu à la fois si central, omnipotent – et multifonctions.

3. L'élaboration des critères de Feighner

En 1972, l'article intitulé *Diagnostic Criteria for Use in Psychiatric Research*, publié dans les *Archives of General Psychiatry* (aujourd'hui intitulé *JAMA Network*) par Feighner, Robins, Guze, Woodruff, Winokur et Munoz (Feighner et al., 1972), proposait une liste de critères s'appuyant sur la méthode

²⁵ Par exemple, dans l'ouvrage de Cahalan, on apprend que l'un des « pseudo-patients » avait été éliminé de l'étude parce qu'il avait déclaré être bien traité pendant son séjour à l'hôpital et avait apprécié l'expérience. Ce patient admet qu'il était d'ailleurs véritablement déprimé à l'époque. De faux détails allant à l'encontre de l'institution psychiatrique avaient été insérés dans le dossier médical d'un autre de ces pseudo-patients. De plus, les pseudo-patients n'avaient pas fait que raconter qu'ils entendaient une unique voix (leur disant « bruit, vide, creux »), mais l'un d'eux (le psychiatre Frank Bartlett) avait dû déclarer qu'il était suicidaire. Cette histoire exemplifie d'ailleurs la « crise de la répliquabilité » en psychologie où de nombreuses découvertes célèbres se sont dissoutes une fois que d'autres chercheurs ont tenté de répéter les expériences.

clinique descriptive. Pour chaque trouble, il existe un ensemble de critères spécifiques (de 2 à 9)²⁶. Les auteurs de cet article, présentant les critères nommés par la suite « critères de Feighner », appartenaient au Département de psychiatrie de l'École de médecine de l'Université de Washington, à Saint-Louis – ensuite nommée « École de Saint-Louis » (Aragona, 2013b). Nous verrons le contexte d'élaboration des critères, leur portée scientifique, leur contenu et les fondations méthodologiques et conceptuelles sur lesquels ils s'appuient, en détaillant dans un premier temps les principes de l'École de Saint-Louis et dans un second temps l'importance des statistiques dans la constitution du DSM.

En ce qui concerne l'élaboration des critères de cet article, le premier auteur, John P. Feighner, affirme s'être appuyé sur les travaux de Sokal et Sneath (Sneath & Sokal, 1973), eux-mêmes s'étant inspirés d'articles hors du domaine de la psychiatrie. Ils présentaient notamment des critères utilisés pour la polyarthrite rhumatoïde, qui peuvent être retrouvés dans la littérature scientifique dès 1957 (Ropes et al., 1957). Ce ne sont pas les données rhumatologiques eux-mêmes qui ont inspiré les auteurs, mais bien la « procédure analytique » que Ropes et collaborateurs ont développé concernant l'élaboration des critères – c'est-à-dire l'intérêt pour l'élaboration de critères fondés sur des mesures statistiques (et notamment, dans l'article, la spécificité et la sensibilité) (Ropes et al., 1957)²⁷. Cet article de 1972 a en tout cas fait date dans l'histoire de la psychiatrie. En termes de portée scientifique, notons qu'en 1989, il était l'article le plus cité de l'histoire de la psychiatrie, comme le rapportaient Horwitz et Wakefield (Horwitz & Wakefield, 2007) et le huitième le plus cité en 2013 avec 4 847 citations (et 7 379 en 2023)²⁸. En ce qui concerne le contenu des critères décrits dans cet article, Feighner et collaborateurs se sont appuyés sur l'histoire de la discipline psychiatrique (Feighner et al., 1972). Cependant, à la différence d'autres spécialités médicales, en psychiatrie, ces données cliniques étaient éparées, hétérogènes, multiniveaux et

²⁶ Les critères de Feighner ne correspondent cependant pas à des critères généraux de définition de ce qu'est un trouble psychiatrique.

²⁷ Nous verrons plus loin que la standardisation de ces critères a été inspirée par des auteurs issus de tradition philosophique comme Hempel (1959).

²⁸ Les articles qui le précèdent sont : celui de l'échelle *Mini-Mental State*, de l'échelle d'Hamilton, de l'échelle HAD de la dépression et anxiété, de l'aspect longitudinal des diagnostics du DSM-III, des critères de recherche que nous discuterons par la suite et de l'échelle PANSS.

fluctuantes dans le temps²⁹. Ces caractéristiques ont conduit les concepteurs du DSM à raffiner la méthodologie initialement proposée. En effet, l'article propose des listes de symptômes appartenant à dix-sept troubles mentaux (par exemple, les « *secondary affective disorders* » [troubles affectifs secondaires] ou les « *obsessive compulsive neurosis* » [névrose obsessionnelle compulsive]) et pour lesquels existaient suffisamment de preuves de fiabilité et de validité à l'époque pour les inclure dans une publication destinée à la recherche. Par exemple, concernant la « névrose obsessionnelle compulsive », à la fois les critères A et B sont requis, avec :

« A. Les manifestations 1 et 2 sont requises : (1) Les obsessions ou les compulsions constituent les symptômes dominants. Ils sont définis comme des idées, des pensées, des images, des sentiments, des impulsions ou des mouvements récurrents ou persistants, qui doivent s'accompagner d'un sentiment de contrainte subjective et d'un désir de résister à l'événement, l'événement étant reconnu par l'individu comme étranger à sa personnalité (c'est-à-dire « égo-étranger ») ; (2) Ils doivent apparaître avant 40 ans. B. Les patients atteints d'un trouble affectif, primaire ou probable, ou atteints de schizophrénie ou de probable schizophrénie, qui manifestent des caractéristiques obsessionnelles-compulsives, ne reçoivent pas le diagnostic complémentaire de névrose obsessionnelle compulsive. Il est particulièrement intéressant de repérer l'algorithme de décision, particulièrement complexe, qui peut être à la fois explicite (par exemple, avec la nécessité des critères A et B) ou implicite (par exemple, les manifestations 1 et 2, ou la présence d'obsessions ou de compulsions) ».

Ils visaient plutôt à « libérer » les chercheurs de la multiplicité des différentes définitions imprécises, ainsi qu'à permettre une recherche plus cumulative, comparable et reproductible – comme l'expliquera *a posteriori* Spitzer dans le cadre de la révision du DSM-III (Spitzer, 1983). L'objectif déclaré était un « terrain d'entente pour différents groupes de recherche » (Spitzer, 1983). Brièvement, nous pourrions

²⁹ Une anecdote désormais rendue célèbre par (Kendler et al., 2010) témoigne que Feighner aurait été inspiré par un auteur nommé Cassidy, ayant publié un article en 1957 portant sur les critères de la dépression. Lorsqu'un autre chercheur parvient à rencontrer ce dernier afin de l'interroger sur les raisons de ces six critères, il aurait obtenu la réponse suivante : « Ça me semblait à peu près correct ». Ce sont ces critères de la dépression, qui « semblaient à peu près corrects » à Cassidy, qui sont aujourd'hui actualisés dans le DSM-5. Nous avons largement critiqué cette anecdote ailleurs, notamment parce qu'elle se veut ironiquement anhistorique, tout en montrant paradoxalement l'importance de la « tradition orale » de la psychiatrie – c'est-à-dire non nosographique.

dire que la fiabilité permet d'assurer la stabilité des décisions cliniques entre experts et dans le temps et qu'une des formes de validité (de construit) concerne la compatibilité avec la structure naturelle des troubles psychiatriques, permettant de comprendre comment les symptômes ont été regroupés au sein des catégories du DSM-III. Les critères de Feighner n'étaient certes pas explicitement formulés pour une utilisation clinique quotidienne.

Mais pour mieux comprendre ces critères, il faut aborder brièvement l'histoire de l'École de Saint-Louis. Nous ne détaillons pas tout cette histoire aux multiples rebondissements, puisque ce groupe de cliniciens s'est rassemblé dès le début des années 1950 (Feighner y est arrivé tardivement, pour effectuer sa « revue de littérature » publiée en 1972) et parce que le détail a été retranscrit par certains des acteurs eux-mêmes (Rodrigo A. Muñoz et George Murphy), notamment dans un article de Kendler (2010) publié dans l'*American Journal of Psychiatry*. Nous pouvons en donner cependant quelques éléments. L'École de Saint-Louis rassemblait au départ les équipes d'Eli Robins de l'Université de Saint Louis au Kansas et de Joseph Zubin au département de biométrie de l'Institut psychiatrique de l'État de New York. Le groupe de travail de l'École de Saint-Louis s'intéressait au diagnostic descriptif et aux bases biologiques des troubles mentaux. Cet intérêt commun est en partie lié à une conception de la psychiatrie comme branche de la médecine (Kirk & Kutchins, 1992) : il s'agissait d'établir une limite entre le normal et le pathologique en décrivant des catégories de troubles mentaux distincts, idéalement témoins d'une cause commune sous-jacente (selon ce qui a été nommé le modèle médical). Pour une part, les membres de l'École de Saint-Louis recherchaient les causes de ces troubles en s'attachant à leurs aspects biologiques et en déterminant les caractéristiques de fiabilité (et idéalement, de validité) selon des méthodes statistiques. Psychométrie et nosologie se rejoignent donc dans le processus de constitution de cette

troisième version du DSM (American Psychiatric Association, 1980)³⁰. Au-delà de ces considérations (psychométriques et statistiques), il est important de préciser combien l'intérêt pour l'approche descriptive, c'est-à-dire l'intérêt pour les symptômes, a été important dans la détermination des catégories : ce ne sont en effet pas les marqueurs biologiques ou les mécanismes physiopathologiques qui permettent de structurer la classification, mais bien les éléments relevés dans la clinique (Wing et al., 2012). En 2014, Stein écrivait ainsi : « Nos connaissances médicales sophistiquées actuelles ont été acquises grâce à des siècles d'observation et de description au cours desquelles les descripteurs et les classificateurs ont joué un rôle aussi dynamique et créatif que ceux qui s'occupent de processus » (Stein, 2014). Ce n'est qu'une fois que ce socle descriptif aurait été établi qu'un projet de nosologie serait réalisable : « Aucune théorie de la maladie ne peut être élaborée avant que le syndrome clinique ait été reconnu et étiqueté », ajoutent ainsi Wing et ses collaborateurs (Wing et al., 2012). À ce socle clinique historique, nous rajoutons avec Steeves Demazeux que ce projet a impliqué à la fois un niveau théorique (le modèle médical, que nous décrirons et discuterons à plusieurs reprises au cours de cette thèse), un niveau méthodologique (des outils spécialisés statistiques et méthodologique, comme la fiabilité ou

³⁰ Par exemple, par la suite, dans le cadre des études sur le terrain (*field-trials*) qui ont été conduites pour les différentes versions du DSM (Clarke et al., 2013) (afin de déterminer la fiabilité des catégories), le coefficient intitulé « Kappa de Cohen » a été majoritairement utilisé (Cohen, 1960). Aboraya, en 2007, a proposé une liste de quinze raisons du manque de fiabilité du DSM : la mauvaise nomenclature psychiatrique, les facteurs liés aux patients (par exemple, anxiété ou problèmes de mémoire), la présentations cliniques du trouble lui-même (potentiellement atypique), l'évolution des symptômes au fil du temps (pour un même patient), la nécessité d'un recours à l'observation du comportement et donc d'un jugement, la dépendance aux proches pour recueillir certaines informations, le style d'entretien du clinicien, son expérience, sa préférence envers certains diagnostics, ses méthodes (ou le manque de méthode) pour structurer son entretien, la formation de sa conception du trouble (dépendant de son « école de pensée »), les contraintes de temps (institutionnelles et financières) ainsi que l'absence d'accord sur les définitions des symptômes psychiatriques. Nous pourrions séparer en trois axes ces raisons de mauvaise fiabilité : celui lié à la nosologie, celui lié au patient et celui lié au clinicien. Les différentes combinaisons de la triangulation de ces trois axes montrent la difficulté à obtenir un diagnostic fiable. Dès 1962, Beck et collaborateurs avaient déjà étudié les raisons du désaccord diagnostique, c'est-à-dire de la faible fiabilité : le facteur nosologique était le plus élevé (62,5 %), les facteurs liés aux patients étaient les plus faible (5 %) et les facteurs liés aux cliniciens étaient de 32,5 % (Beck et al., 1962). Dans l'étude d'Aboyara de 2007, 45 ans plus tard, la déception liée à la nosologie est moins importante : les cliniciens semblaient plus satisfaits de la nosologie (14,9 % de déception pour la fiabilité) que de leur propre pratique (expliquant 63,5 % du manque de fiabilité), pour 21,6 % du manque de fiabilité lié au patient lui-même (Aboraya, 2007). Les cliniciens témoignent donc du rôle majeur du DSM dans la suppression des lacunes de l'ère pré-DSM, ou des anciennes versions du DSM, du moins en ce qui concerne la clarté nosologique.

l'opérationnalisation) et un niveau institutionnel (c'est-à-dire un front commun de recherche) (Demazeux, 2013a).

Pour résumer, la constitution du DSM-III s'est en partie adossée sur des listes de critères descriptifs spécifiques à chaque trouble psychiatrique, correspondant aux critères de Feighner, qui ont été préparés pour dix-sept troubles psychiatriques par la suite intégrés au DSM-III. L'élaboration de ces critères a été guidée par la volonté d'améliorer la fiabilité de la psychiatrie, entraînant, comme nous le verrons, une nécessaire structuration de la discipline face aux multiples enjeux à laquelle elle était soumise.

4. L'élaboration des critères diagnostiques pour la recherche

En parallèle des critères de Feighner et des groupes de travail (et en s'appuyant sur ces travaux), en 1978 et à la demande de l'Institut national de la santé mentale (*National Institute of Mental Health – NIMH*), Spitzer et collaborateurs ont développé les critères diagnostiques pour la recherche (*Research Diagnostic Criteria – RDC*) afin de tenter d'améliorer la (faible) fiabilité des diagnostics psychiatriques (Spitzer et al., 1978). Les RDC ont montré de bons résultats en termes de fiabilité. Par exemple, pour les troubles dépressifs caractérisés, la fiabilité initiale était de 0,97 (Spitzer et al., 1975). À l'image des critères de Feighner développés pour dix-sept troubles, les RDC proposaient des critères pour vingt-cinq troubles et pour de nombreux autres sous-types de troubles (Spitzer et al., 1978). Dans cette section, nous décrirons notamment pourquoi ces critères pour la recherche ne sont pas des « critères pour la clinique » (ce point sera notamment au cœur des discussions sur les RDoC). Nous décrirons également les caractéristiques des RDC.

Les RDC introduisent la notion de critères « monothétiques » et « polythétiques ». Au sein de chaque trouble psychiatrique, il existe : i) un ou plusieurs critères monothétiques, dont la présence d'au moins l'un d'eux est nécessaire pour poser le diagnostic de trouble ; ii) et plusieurs critères polythétiques,

dénommes ainsi car plusieurs d'entre eux sont nécessaires pour poser le diagnostic. Par exemple, pour le trouble dépressif caractérisé, il faut (au moins) un de ces deux critères monothétiques : anhédonie (une diminution marquée du plaisir pour les activités) *ou* tristesse de l'humeur ; *et* au moins quatre critères polythétiques parmi les sept suivants : perte ou gain de poids significatif, insomnie ou hypersomnie, agitation ou ralentissement psychomoteur, fatigue ou perte d'énergie, sentiment de dévalorisation ou de culpabilité, diminution de l'aptitude à penser ou à se concentrer et / ou pensées de mort récurrentes. Cette distinction entre critères monothétiques et polythétiques a été introduite par Samuel Guze (appartenant à l'École de Saint-Louis) (Guze, 1978) et l'idée provient également d'autres classifications médicales (comme celle de la classification de la polyarthrite rhumatoïde – (Ropes et al., 1957)³¹). Guze est parti du principe qu'il n'existait pas de caractéristique suffisamment spécifique, en psychiatrie, pour affirmer la présence d'un trouble sur la base d'un seul critère (Guze, 1978). Une telle distinction entre des critères monothétiques « nécessaires » pour affirmer la présence du trouble et des critères polythétiques (non nécessaires ni suffisants) semble importante : elle permet de proposer une catégorie diagnostique utile en pratique, tout en acceptant que les troubles psychiatriques ne puissent être entièrement réifiés, implicitement, en étant réduits à des ensembles fixes de critères immuables. Par exemple, la combinaison des différents critères monothétiques et polythétiques peut aboutir à 1 030 profils de symptômes uniques de dépression, permettant de relativiser l'importance des critères pour définir ce qu'est théoriquement un trouble – pour la plupart du temps dépendant des tableaux cliniques rencontrés en pratique clinique³².

³¹ Pour revenir à cette comparaison avec la rhumatologie, plusieurs auteurs ont justement relevé les ressemblances entre la pathologie psychiatrique en général et l'atteinte rhumatismale qu'est la polyarthrite rhumatoïde : leur étiologie est inconnue (bien qu'un certain nombre de marqueurs biologiques aient été découverts depuis pour la seconde, tel que le facteur rhumatoïde), « les signes les plus fréquents et les plus apparents ne sont pas les plus spécifiques » et les désaccords restent fréquents dans la pratique diagnostique (par exemple, actuellement, le débat s'est déplacé, en médecine interne, vers l'exploration de la différence entre spectre inflammatoire et spectre auto-immun) (Demazeux, 2013b). Notons cependant qu'on retrouve déjà en 1944 cette distinction dans la classification de Jones en rhumatologie (Jones, 1944).

³² Il est classiquement admis qu'un diagnostic en psychiatrie devrait faire des prédictions sur l'étiologie, l'apparition de la maladie, son évolution et ses récurrences et ainsi permettre le développement de stratégies de prévention et de traitement efficaces. Il semblerait que les caractéristiques importantes d'un diagnostic utile soient : un nombre limité de symptômes spécifiques à la maladie, ainsi qu'un nombre limité de facteurs de risque, comportant un *pool* homogène de symptômes, offrant une étiologie clairement circonscrite facilitant la détection précoce et la prévention. Or, en prenant l'exemple de la dépression, celle-ci est très hétérogène sur les deux

III. Les enjeux du DSM-III : entre opérationnalisation et contingence

Dans cette section, nous soulignerons l'importance de l'opérationnalisation dans le cadre du développement du DSM-III. Nous verrons que cette notion est fortement liée aux enjeux de fiabilité précédemment décrits. Le principal objectif de présenter ces notions est de bien comprendre pourquoi l'approche par catégorie, qui repose sur la volonté de rendre plus manipulable les entités psychiatriques, semblait nécessaire dans le contexte de l'époque. Cette volonté était en fait rendue nécessaire par les multiples problématiques que devaient relever le DSM-III naissant. Au-delà d'une volonté d'amélioration de la classification, un certain nombre de facteurs contingent ont contribué à promouvoir celle-ci.

1. Importance de l'opérationnalisation

Tout en reconnaissant qu'un système fiable n'est pas nécessairement valide (la fiabilité pourrait simplement correspondre au fait que tout le monde obtient la même mauvaise réponse), Spitzer a souligné qu'une bonne *validité* requerrait, avant tout et minimalement, une bonne *fiabilité*. Même s'il manque de validité, un système nosologique fiable pourrait fournir un point de départ scientifiquement adéquat à partir duquel les chercheurs développeraient un système plus valide (Spitzer, 1983). Nous avons vu que les critères de Feighner de 1972 sont mieux formalisés et systématisés par les RDC de 1978

dimensions des symptômes et de l'étiologie : elle est associée à un grand nombre de symptômes et d'innombrables voies peuvent conduire à un épisode dépressif caractérisé.

et sont ultimement institutionnalisés par leur intégration au sein du DSM-III, en 1980, manuel dont la coordination générale revint à Spitzer (on notera qu'il n'est pas un co-auteur de l'article de 1972 présentant les critères de Feighner). Comme nous l'avons souligné, ces ensembles de critères permettent d'éviter de quelconques inférences sur la causalité (Wilson, 1993), forme de neutralité qui sera conservée par la suite lors des différentes versions du DSM. Un des principaux objectifs des deux principales initiatives reconnues comme fondatrices du DSM (à savoir justement les critères de Feighner et les RDC) était de créer un système qui classait les troubles psychiatriques : i) selon des critères opérationnalisés, ii) dont les catégories comporteraient une bonne fiabilité (en espérant qu'elle conduise à terme à une bonne validité), iii) sans que ces critères ne dépendent de perspectives théoriques (et notamment, psychodynamiques).

Il est ici important de s'arrêter sur la définition du terme « opérationnel » (Hempel, 1965). En effet, ce point est central pour notre problématique : en comprenant les enjeux (d'opérationnalisation) en psychiatrie, nous saisissons toute l'ampleur du projet de l'époque, qui veut à la fois légitimer les catégories et mieux saisir et ancrer la sémiologie dans la pratique. Le terme d'opérationnalisation trouve son origine dans la physique moderne (Bridgman, 1927), dans laquelle la définition de « l'essence » d'un objet a été remplacée par la description des opérations de mesure dans un contexte expérimental. L'opérationnalisation considère que la *signification des concepts est déterminée par l'ensemble des opérations utilisées pour leur mesure*. Bridgman écrivait ainsi que l'opérationnalisation se confond (« est synonyme de ») la démonstration de cet ensemble d'opérations (Bridgman, 1927). Appliqué à la psychiatrie, le terme « opérationnel » est exigeant : la description d'*opérations* par des preuves expérimentales peut se révéler particulièrement complexe et ambiguë. Il s'agit d'ailleurs d'une objection classique à l'opérationnalisme : le concept d'opération est vague car la notion même de ce qu'est une opération l'est tout autant (Chang, 2021). Alors qu'est-ce que pourrait être une opération en psychiatrie clinique ? Comme nous l'étudierons par la suite en nous penchant sur la question de la validité (de construit) (Cronbach & Meehl, 1955), il ne peut y avoir d'opération en psychiatrie que sur des objets

(comme des critères) empiriquement manipulables. La question qui se pose alors est de savoir si les signes et les symptômes de la psychiatrie constituent des entités empiriquement manipulables, ou ne sont justement, eux-mêmes, que des « construits », c'est-à-dire des réseaux de propriétés ne pouvant être approchés directement par l'expérience. En somme, dans cette partie, nous admettrons que ce qui fait une opération en psychiatrie soulève des problématiques à la fois nosologiques, concernant les construits que sont les catégories psychiatriques, tout en interrogeant la sémiologie. Malgré ces exigences pratiques, il semble que les travaux des chercheurs et cliniciens des années 1970 ont permis de garantir une forme (dégradée³³) d'opérationnalisation pour les critères nosographiques utilisés en psychiatrie. L'opérationnalisation permettrait d'améliorer la fiabilité et l'harmonisation des diagnostics et des soins (Spitzer et al., 1980). Feinstein écrivait :

« La production de critères opérationnels a constitué une avancée pionnière et unique en nosologie... Dans le domaine de la nosologie diagnostique, l'établissement de critères opérationnels représente une avancée aussi évidente, nécessaire, fondamentale et importante que le progrès obtenu en obstétrique et en chirurgie lorsque Semmelweis, Oliver Wendell Holmes et, plus tard, Lord Lister, ont exigé que les obstétriciens et les chirurgiens se lavent les mains avant d'intervenir sur le corps humain » (Feinstein, 1970).

Cette comparaison apparemment étonnante souligne que les critères nosographiques de la psychiatrie sont présentés comme un préalable au soin, dans le souci de ne pas être iatrogène au travers des théories portées par les praticiens et de leur subjectivité. Les orientations théoriques divergentes des années 1970 constituent en effet une des pierres d'achoppement de la psychiatrie, sur lesquels les concepteurs du

³³ Nous utilisons ce terme pour signifier que l'opérationnalisation en psychiatrie est sans aucun doute différente de celle développée en physique, bien que ce constat concernant la physique nécessiterait d'être nuancé étant donné la complexité (parfois non-apparente) de celle-ci – (Ruphy, 2016).

DSM-III se sont particulièrement penchés. Ils ont cherché à rendre plus acceptable le manuel en le dégageant de la subjectivité des cliniciens et chercheurs individuels³⁴.

2. Origine de l'opérationnalisation

Kendler et Parnas ont fourni une analyse détaillée de l'importance des travaux philosophiques d'Hempel dans le processus d'élaboration du DSM-III (Kendler & Parnas, 2012). Ils insistent sur le fait que les premiers auteurs du manuel, dont Spitzer, « n'en étaient pas les premiers concepteurs ». Certes, ces auteurs se réclament de la pensée « opérationnalisante » des critères de Feighner de 1972 (Feighner et al., 1972). Mais peu admettent avoir été influencés par Hempel, philosophe des sciences, qui en 1959 est invité par l'*American Psychopathological Association* (qui est différente de l'*American Psychiatric Association* (APA) et de l'*American Psychological Association*) pour prononcer un discours sur la nosologie psychiatrique. Depuis au moins son article de 1935 intitulé « *L'analyse logique de la psychologie* », Hempel a proposé l'utilisation de critères opérationnels pour la science. De tels critères permettraient d'obtenir une classification plus *scientifique* pour les troubles psychiatriques. Hempel soulignait que les deux principales exigences d'une classification prétendument scientifique étaient : i) l'utilisation d'un vocabulaire scientifique (contenant des termes descriptifs et / ou des termes théoriques) ; et ii) l'utilisation de critères opérationnels. Hempel décrit également la manière dont une discipline (en l'occurrence, ici, la psychiatrie) se construisait en deux mouvements : i) un mouvement historique

³⁴ Plus précisément, dix objectifs de cette opérationnalisation ont pu être développés : i) perfectionner l'utilité clinique du DSM, afin de prendre des décisions de traitement et de gestion dans divers contextes cliniques, ii) améliorer la fiabilité des catégories diagnostiques, iii) rendre plus acceptable le manuel pour des cliniciens et des chercheurs d'orientations théoriques divergentes, iv) être utile pour la formation des professionnels de santé, v) maintenir la compatibilité avec les codes de la CIM (dans sa neuvième édition à l'époque), vi) éviter une nouvelle terminologie et de nouveaux concepts qui rompraient avec la tradition, vii) tenter de parvenir à un consensus sur la signification des termes diagnostiques nécessaires, viii) pouvoir servir en clinique comme en recherche, ix) tout en restant modulable et critiquable, x) et comme nous le développerons succinctement, rendre cohérente la classification avec les données de la recherche portant sur la validité des catégories diagnostiques (Beals et al., 2004).

décrivant la symptomatologie et privilégiant l'observation, qui ferait partie du « siècle précédent » en ce qui concerne la psychiatrie ; ii) une étape théorique lors de laquelle était décrite l'étiologie des troubles, qui devait s'élaborer sur la base d'un langage fiable et sans considération appréciative (Hempel, 1965; Kendler et al., 2010). Cette influence d'Hempel sur l'élaboration du DSM (comme ensuite de la CIM) a été soulignée dans un article datant de la même année (1959) d'Erwin Stengel, comme nous l'avons dit, chargé par l'OMS d'établir des recommandations pour une classification la plus largement acceptable (Stengel, 1959). Si comme le rapporte Bentall dans *Madness explained: Psychosis and human nature*, Mishara et Schwartz en 2013 ou Bolton dans *Classification and causal mechanisms: A deflationary approach to the classification problem* (2012), les concepteurs du DSM n'ont pas admis avoir eu connaissance du texte de conférence de Hempel, ses idées auraient pu être profondément intégrées dans le style de pensée des scientifiques de l'époque – selon une acculturation rendue possible par le contexte positiviste et la promesse de standardisation uniforme des dénominations de la psychopathologie. Cependant, un certain nombre de recherches récentes tempèrent fondamentalement l'influence de Hempel sur le DSM, tel que l'illustre le titre d'un article de Cooper et Blashfield de 2018 : « *The myth of Hempel and the DSM-III* » (Cooper & Blashfield, 2018)³⁵.

3. Contextualisation socio-historique du DSM-III

Les multiples origines du DSM-III permettent de comprendre que cette version de la classification est aussi le produit final d'un ensemble de contingences historiques et contextuelles. Par exemple, le « déclin » de la psychanalyse avait pu être imputé à la parution du DSM-III, même si aujourd'hui un certain nombre d'historiens ont montré que sa perte de vitesse était antérieure à sa publication (Hess &

³⁵ Différents arguments ont pu être avancés à l'encontre de la prééminence de Hempel dans la constitution du DSM et notamment Aragona, qui a examiné en 2013 la relation historique et épistémologique entre la philosophie néopositiviste des sciences et le DSM-III (Aragona, 2013a).

Majerus, 2011; Shorter, 2008). Nous utiliserons trois exemples illustrant la contingence du développement du DSM et de ces critères. Pour le premier exemple, il s'agit de l'influence de la découverte d'un thérapeutique médicamenteuse sur le développement des critères ; pour le deuxième, il s'agit de souligner l'importance de la conduite des entretiens telle qu'elle était réalisée antérieurement au DSM-III dans les enjeux de fiabilité ; pour le troisième exemple, nous soulignons l'importance de facteurs sociologiques dans la constitution du DSM-III³⁶.

Premièrement, en prenant l'exemple du trouble bipolaire, Hacking a pu décrire que la découverte de nouveaux traitements avait largement contribué à l'ancrage du DSM-III dans les pratiques de soin (Hacking, 2013). Il cite la découverte d'un traitement efficace pour contrôler la manie : le lithium (Cade, 1949). Avec un certain scepticisme du fait des potentiels dangers des effets secondaires du lithium, la *Federal Drug Administration* (FDA – l'agence de sécurité des médicaments américaine) a approuvé son utilisation médicale en 1970 et son utilisation dans la prise en charge de la dépression maniaque en 1974. Or, comme nous l'avons souligné, en 1972, paraissait les premières descriptions opérationnalisées des critères de Feighner. La possibilité de se fier à des catégories de troubles stables en assurant un certain accord entre les cliniciens permettait une application standardisée d'un tel traitement. En effet, en adaptant les indications du traitement à des critères fiables, la sécurité du lithium pouvait être théoriquement contrôlée (Hacking, 2013). Ainsi, en partie et indirectement, l'arrivée de nouvelles thérapeutiques a soutenu le développement et l'instauration dans les pratiques des critères d'opérationnalisation.

Deuxièmement, bien avant les années 1970, le développement des entretiens structurés (ou standardisés) avait été motivé par la nécessité d'améliorer la fiabilité des évaluations psychiatriques. Le projet de

³⁶ Nous tenons à exprimer notre volonté d'être prudent à ce stade. En effectuant un travail relativement biaisé de reconstitution de l'histoire (nécessairement contingente) d'une classification au travers de trois exemples, nous souhaitons simplement montrer que la paternité du DSM-III et son influence ne doivent pas être imputées à une succession de découvertes et publications scientifiques, mais qu'elles devraient être replacées dans le contexte général de l'époque, comportant de multiples facteurs contingents scientifiques, institutionnels, politiques et financiers.

diagnostic américano-britannique parrainé par l'OMS (Gibson, 1973) a démontré qu'il existait des « habitudes diagnostiques » nettement différentes chez les cliniciens britanniques et américains. Il ressortait clairement de ces études qu'une « science psychiatrique » ne serait possible qu'en renforçant la fiabilité des évaluations psychiatriques (Gibson, 1973). Ce projet, associé à celui de l'*Epidemiologic Catchment Area* de la fin des années 1970 (et également à l'*US Mental Hygiene Movement*, à la *Midtown Manhattan study* ou aux *Stirling County studies*, que nous ne détaillerons pas plus – (Coleman, 1988; Srole & Fischer, 1978)), a également démontré que les différences diagnostiques (et notamment la fiabilité) pouvaient être minimisées en utilisant des entretiens structurés et standardisés, standardisation offerte par une critériologie comme celle de Feighner et des RDC. La conduite de ce genre d'études de grande ampleur a ainsi pu conduire à une volonté d'intégrer des critères plus opérationnels.

Troisièmement, afin d'illustrer comment l'intégration de facteurs sociologiques ont pu jouer un rôle dans le processus d'élaboration du DSM-III, nous détaillons brièvement le processus d'intégration des critères de la schizophrénie au sein du DSM-III. L'approche schneiderienne du diagnostic de la schizophrénie est née en partie des écrits de Kurt Schneider et Mayer-Gross, qui ont été rassemblés dans le *Clinical Psychopathology*, publié en allemand en 1950 et traduit en anglais en 1959 (Schneider, 1959). Quelques années plus tard, en 1961, John Wing développait une évaluation des symptômes positifs de la schizophrénie reposant largement sur les symptômes décrits par Schneider et Mayer-Gross (Wing, 1961). Lorsque Spitzer, Endicott et Robins ont été chargés d'élaborer les RDC, ils ont utilisé les travaux de Wing qui étaient d'une grande fiabilité (Spitzer et al., 1975). C'est en rapport avec cette description initiale de Schneider et Mayer-Gross, puis sa traduction anglophone tardive, sa redécouverte par Wing et sa relecture par Spitzer et collaborateurs que les critères diagnostiques ont été intégrés dans le DSM-III. Ainsi, dans les différentes versions du DSM après 1980, on retrouve la description schneiderienne des différentes formes d'hallucinations auditives, du « commentaire » à la « voix » (Mellor, 1970). La chaîne d'événements historiques qui a conduit à l'inclusion de l'approche schneiderienne dans le DSM-III aurait pu être interrompue « à toutes ses étapes » (Kendler, 2009) si les différents protagonistes ne

s'étaient pas influencés, impressionnés, ou n'avaient pas été soumis aux aléas historiques (traductions, relectures, republications, etc.) que nous venons de décrire.

Ces trois exemples permettent d'illustrer la contingence qui rythme la dynamique de toute discipline scientifique et médicale et notamment de la nosographie psychiatrique. La thèse de ce chapitre est supportée par l'existence de telles contingences, associées aux enjeux de fiabilité et d'opérationnalisation précédemment décrits : pour répondre à une multiplicité de fonctions pour la psychiatrie, les cliniciens travaillant à la préparation d'une nouvelle classification ont dû proposer et consolider des catégories pour la nosologie³⁷.

IV. L'évolution des versions du DSM après le DSM-III

Nous détaillons désormais le développement des versions ultérieures au DSM-III, sans chercher à retracer entièrement l'origine, le contexte ou les pressions en jeu pour chacune des versions. Une littérature abondante sur les potentielles révisions et processus d'élaboration du DSM-5 a déjà été fournie

³⁷ Plus précisément, les cliniciens ont tenté de répondre à l'absence « d'agrégateur » causal sous-jacent (ou variable latente) en organisant la classification sur les avis d'expert (fiabilité) et l'opérationnalisation (American Psychiatric Association, 1980, 2013 ; Feighner, 1972 ; Demazeux, 2013). Certes, il est vrai que certains symptômes sont *plus fréquemment associés*. Par exemple, une élévation de l'humeur est plus susceptible d'être accompagnée d'une agitation psychomotrice que d'une anhédonie. Mais en pratique courante, les symptômes ne s'organisent pas en grappes nettes, ils ne possèdent pas de limites distinctes (comme c'est le cas dans les catégories du DSM, ou même dans la HiTOP). La vision catégorielle de la dépression en tant que trouble discret n'est pas cohérente avec les données taxométriques ou psychométriques (Kendell, 1989; Markon et al., 2011; Ruscio et al., 2009). Les symptômes de dépression dans les échantillons en population générale forment des distributions qui ne se chevauchent pas entre individus sains et déprimés. De plus, comme cela est fréquent en clinique et comme Beard et collaborateurs l'ont montré dans une analyse en réseau des symptômes de dépression et d'anxiété (Beard et al., 2016), l'élévation de l'humeur peut également être fréquemment associée à des pensées inquiétantes et conduire à une catatonie – des symptômes d'allure « négative » qui sont elles-mêmes fréquemment associées à une humeur basse. Cette difficulté à délimiter de manière nette les catégories est renforcée par la présence de symptômes ubiquitaires, tels que les problèmes de sommeil (Blanken et al., 2020). Enfin, le caractère polythétique des regroupements syndromiques (qui fait que les individus n'ont besoin de répondre qu'à un sous-ensemble de critères pour être diagnostiqués) démultiplie les formes symptomatiques d'une catégorie, témoignant d'une variabilité considérable entre des individus ayant le même diagnostic.

(par exemple, consulter *Making the DSM-5* (Paris & Philipps, 2013), *The DSM-5 in Perspective* – (Demazeux & Singy, 2015) ou (Kendler, 2013)). Dans l’appréciation de l’histoire et de la portée épistémique des différentes versions du DSM (comme plus généralement dans toute étude historique), il est décisif de séparer les conditions de la création de l’objet d’étude, les enjeux qu’il introduit et les usages qui en ont été fait. Basé sur les section précédentes, nous résumons ici les conditions de la création des versions suivantes du DSM (IV et 5 et leurs révisions), puis nous aborderons brièvement ses enjeux et ses usages.

1. L’élaboration du DSM-IV

Le DSM-III-R a été publié en 1987, le DSM-IV en 1994, le DSM-IV-TR en 2000, le DSM-5 en 2013 et le DSM-5-TR en 2022. Ces éditions ultérieures au DSM-III-TR possèdent de nombreuses caractéristiques semblables à celui-ci et notamment une critériologie mono- et polythétique et un fort intérêt pour la fiabilité. Nous pouvons relever que le nombre de catégories de troubles est passé de 265 troubles dans le DSM-III, à 292 dans le DSM-III-R et à environ 297 dans le DSM-IV et DSM-5³⁸ (le relevé est nécessairement imprécis en raison des recoupements et de la manière dont on « compte » les troubles ou les catégories) (Pressman, 1997). Les ajouts effectués lors des publications successives du DSM ont également inclus des informations recueillies à partir d’études sur la physiopathologie des troubles mentaux et, plus récemment, des données obtenues grâce à des études de neuroimagerie (Kawa & Giordano, 2012). Ainsi, chacune des versions a apporté d’importants changements au contenu des

³⁸ Il est également important de ne pas confondre cette inflation diagnostique avec la potentielle augmentation de troubles diagnostiqués. En effet, Wakefield a soutenu (dans une discussion avec Houts ayant duré d’au moins 1996 à 2001 – (Houts, 2001)) que la « thèse de la discontinuité » n’était pas tenable en psychiatrie : la plupart des troubles qui sont diagnostiqués avec les dernières versions du DSM (le DSM-IV dans le cas de Wakefield) *l’auraient été* en utilisant le DSM-III (mais bien entendu, selon d’autres critères et intégrés à d’autres catégories) (Wakefield, 2001). Autrement dit, il ne faut pas confondre l’inflation des catégories diagnostiques et la baisse du seuil du normal et du pathologique.

catégories et au nombre de troubles, mais il faut admettre sans trop de difficulté qu'aucune modification substantielle dans le style général ou la critériologie n'a été adoptée depuis le DSM-III (Jablensky & Kendell, 2002). Il est cependant important de noter que c'est à l'occasion de la parution du DSM-IV, en 1994, que l'inclusion des critères descriptifs dans la dixième révision de la CIM a été réalisée. Comme nous l'avons dit, l'OMS avait déjà discuté de la possibilité d'inclure des listes de critères diagnostiques dans le chapitre V de la CIM-9, publiée en 1975, parallèlement aux descriptions cliniques narratives – sans pourtant que cette possibilité ne voit le jour. Cette intégration n'a donc été réalisée que pour la CIM-10, publiée en 1992 (World Health Organization, 1990). La production de cette version (conjointement à la construction du DSM-IV) a pris environ 10 ans, de 1982 à 1992 (Üstün & Sartorius, 1995), nécessitant le développement de nombreuses études de terrain (Robins et al., 1989), d'instruments de psychométrie et de plusieurs groupes de travail (neuf exactement – et rassemblant environ 200 psychiatres lors de la Conférence de 1982 sur le sujet).

2. L'élaboration du DSM-5 : un processus conservateur

Avant sa parution, le DSM-5 avait promis de grands changements. Au début de son élaboration, en 2002 et sous la coordination de David Kupfer, il ne tendait pas seulement à assouplir l'appréciation des catégories en *dimensionnalisant* les entités de l'Axe I du DSM-IV-TR : il s'est petit à petit orienté, entre 2002 et 2008, vers une véritable réflexion concernant un changement radical de conception des troubles psychiatriques (First, 2004; Kupfer, 2005). Il s'agissait de modifier l'appréhension de la psychopathologie en proposant une pluralité de combinaisons décrivant « l'étendue de la psychopathologie humaine » (Widiger & Clark, 2000) – cité par (Demazeux, 2013a). Pour cela, du moins au début du projet du DSM-5, la méthode adoptée se basait sur une évaluation en clusters de la validité

dimensionnelle, cette validité étant le témoin de nouveaux outils diagnostiques et thérapeutiques. Mais dès 2002, Jablensky et Kendell prévoyaient que le DSM-5 se heurterait à des limites méthodologiques :

« Les difficultés avec les modèles dimensionnels de psychopathologie proviennent de leur nouveauté : manque d'accord sur le nombre et la nature des dimensions nécessaires pour rendre compte de manière adéquate des variations cliniquement pertinentes ; absence d'une mesure établie et empiriquement fondée pour évaluer la gravité ou le changement ; et, peut-être plus important encore, une complexité et une lourdeur de ces modèles dans la pratique clinique quotidienne » (Jablensky & Kendell, 2002).

C'est finalement un texte relativement « conservateur » qui est approuvé en 2012 (Demazeux & Singy, 2015)³⁹. Ce conservatisme, lié pour une grande part à l'enracinement du DSM dans son environnement clinique et de recherche, témoigne cependant d'une force de stabilité : ses ambitions et son projet sont restés globalement similaires au fil des révisions des trente dernières années. Aragona écrivait :

« Le DSM-III néo-kraepelinien et le DSM-5 sont tous deux fondés sur la même hypothèse de validité : les descriptions cliniques provisoires [...] seront pleinement validées lorsque certains facteurs neurobiologiques spécifiques seront 'découverts' confirmant que la description clinique correspond bien à une maladie neurobiologique. [...] La direction de l'entreprise de découverte [va] du syndrome à la physiopathologie sous-jacente dans les DSM » (Aragona et al., 2014).

La notion de « dépendance au chemin », qui s'est répandue dans la littérature scientifique notamment à la suite d'un article de Liebowitz et Margolis, *The Fable of the Keys* (Liebowitz & Margolis, 1990) (**Annexe 1 – « Dépendance au chemin et régime des promesses dans la nosologie psychiatrique »**), permet de rendre compte de quelques-unes des raisons de cette inertie (Garud et al., 2010). Le processus

³⁹ Sans même plus entrer dans le détail, il est important de souligner que la prise en compte des facteurs contextuels et des déficiences et handicaps est réalisée par un système relégué à la fin de l'ouvrage. Il en est de même pour les troubles de la personnalité, dont le système dimensionnel est simplement proposé comme piste de recherche future.

de dépendance au chemin explique en partie pourquoi la révision du DSM est devenue si difficile (Demazeux & Singy, 2015) : le « verrouillage » d'une classification correspond au fait que ses modifications (et *a fortiori* son remplacement) est d'autant plus difficile que son ancrage social, politique, clinique ou économique est important. Le DSM constitue une composante à part entière de l'infrastructure scientifique de la santé mentale. En effet, il a pu garantir une bonne fiabilité à la psychiatrie (apport clinique) et à la recherche (apport scientifique), il a créé un « pare-feu » entre la psychiatrie et la psychanalyse à l'époque de la constitution du DSM-III, a proposé un *livre de codes* déterminant les remboursements (apports administratif et économique), a fourni à la FDA un système d'approbation des médicaments pour mener des essais cliniques (apports industriel et législatif) et a donné un air de légitimité à la psychiatrie en l'intégrant plus fermement dans le paysage médical. Il a de plus soutenu les lignes directrices des programmes de santé mentale ou des algorithmes hospitaliers et a constitué un soutien institutionnel à l'APA, qui a ainsi acquis (pour le meilleur et pour le pire) une position de centralité scientifique et administrative en santé mentale. Une telle infrastructure scientifique a donc un fort ancrage dans différents champs de la société et son processus itératif de révision en est d'autant plus délicat⁴⁰. Ainsi, en décrivant la relative stabilité du DSM au fil de ses versions, nous

⁴⁰ La conservation du modèle du DSM au fil du temps a été en grande partie assurée (et assumée) par l'espoir conféré par le principe d'itération épistémique, proposant qu'un raffinement progressif de la classification serait à l'œuvre tout au long de son déploiement. Ce raffinement, nommé « itération épistémique », a donc en partie légitimé la conservation du DSM, tant qu'il n'aurait pas « imposé » (selon les termes de Chang), afin d'éviter tout retournement excessif, selon un principe de conservatisme décrit par Kendler (Kendler, 2013). La notion d'itération épistémique semble en effet particulièrement importante dès lors qu'on considère les classifications. Elle permet d'endosser l'imperfection liée à l'idéalisation d'une classification, tout en explicitant ses limites potentielles, les systèmes alternatifs parallèles susceptibles de lever ces limites et les éventuelles divisions (à venir) du domaine. Elle assume l'existence d'un modèle imparfait, mais qui ne pourra que s'améliorer en se confrontant aux données. Cependant, en cherchant constamment à corriger ses imperfections (par exemple, son manque de correspondance avec les phénomènes cliniques ou neuroscientifiques), les progrès de la classification peuvent la conduire à tomber dans un « mauvais minimum local » (Chang, 2004). Ce mauvais minimum local correspond à l'idéalisation certes la plus raffinée possible, mais dépendante du chemin parcouru (Liebowitz & Margolis, 2014). En retour, l'itération épistémique propose des solutions pour éviter à une classification de tomber dans cette dépendance au chemin aboutissant à un mauvais minimum local. Par exemple, une de ces solutions correspond à la possibilité de développer un pluralisme (empirique) pour augmenter la diversité des orientations du progrès scientifique. La tension vers l'idéalisation peut donc conduire à développer une pluralité de classifications, comme nous les décrivons ici. Nous soutenons ainsi que grâce à l'idéalisation explicite et assumée d'un modèle de classification comme le DSM, une certaine forme d'*humilité* a permis la conservation du modèle au prisme de la notion d'itération épistémique. Cette notion d'itération épistémique engage à une

comprenons d'autant mieux son ancrage dans le paysage de son époque et avec lui les multiples fonctions qu'il est censé remplir.

3. Implications de l'ancrage institutionnel du DSM en psychiatrie

Lors de la rédaction du DSM-III, le défi consistait à justifier le fait que la psychiatrie pouvait être une discipline appartenant légitimement à la médecine et non un simple instrument de contrôle social⁴¹. Le DSM-III a subi de nombreuses pressions, il a été critiqué (autant qu'il a été soutenu). De fait, son histoire relève bien plus du « tour de force » que de la chance ou du hasard (Demazeux, 2013a). Blashfield et collaborateurs (dans *The Cycle of Classification: DSM-I through DSM-5*) décrivaient la manière dont cette réussite s'est « gagnée au fil du temps, par des séries de publications scientifiques et de convictions fortes ». Les luttes d'influence, qui sont encore bien visibles aujourd'hui, étaient d'ailleurs déjà présentes lors de son lancement (Decker, 2013). Comme nous l'avons vu, les critères de Feighner ont su s'intégrer dans l'environnement (favorable) dans lesquels ils se sont déployés. Les raisons de ce succès sont également liées au fonctionnement de la recherche et de la productivité scientifique, dépendant d'un « effet Matthieu », tel que décrit par Merton (l'application à la psychiatrie est reprise à Steeves Demazeux) : dans les disciplines scientifiques, plus on réussit (en termes de publication, d'influence, de levée de fonds), plus on réussit facilement (Blashfield et al., 1992; Merton, 1968). Cette réussite peut prendre de l'ampleur dans un cercle restreint de personnalités influentes, par exemple en lien avec la loi de Lotka de la productivité scientifique : en nombre relatif, peu de chercheurs publient par rapport au nombre total

diversité d'approches, ce qui constitue une explication au pluralisme nosologique actuellement observé en psychiatrie.

⁴¹ Cependant, comme l'ont noté un certain nombre d'auteurs, des désillusions sont survenues assez rapidement dans l'histoire des dernières versions du DSM, comme l'illustre par exemple un ouvrage de Whooley et Horwitz de 2013 intitulé : *The Paradox of Professional Success: Grand Ambition, Furious Resistance, and the Derailment of the DSM-5 Revision Process* (Whooley & Horwitz, 2013).

de personnes concernées par une discipline en particulier (Yablonsky, 1980). Ainsi, une majorité relative de petite taille constituée d'auteurs peut influencer l'intégralité d'un champ scientifique.

Les implications du DSM ont été nombreuses. Et parmi elles, de nombreuses conséquences des révisions du DSM n'ont pu être anticipées. En 2007, Andreasen rappelait ces imprévus qui ont fait suite à la parution du DSM-III, comme l'appauvrissement de la formation clinique des psychiatres ou « l'effet déshumanisant » d'une pratique qui propose de relever des critères sans prendre en compte l'histoire de vie de la personne (Andreasen, 2007). Spitzer connaissait l'importance de la « prévention d'amont » et avait cherché à anticiper ces critiques : il avait précisé que le DSM ne devait être utilisé que par les cliniciens (et non pas hors de la clinique, au niveau social, des assurances, de la justice, etc.) et il n'en faisait qu'une étape dans la compréhension globale des troubles psychiatriques (Spitzer, 1983). Il rappelait que de nombreux critères n'étaient issus que d'un consensus d'experts et que l'histoire clinique des patients devait impérativement être prise en compte dans le même temps que le relevé des critères (Spitzer et al., 1978)⁴². Pour toutes ces raisons, le DSM pourrait avoir « échappé à ses concepteurs » (Demazeux, 2013a). Les conséquences du succès du DSM-III, pour la plupart inattendues, n'avaient en fait pas pu être prévues. Plus exactement, le manuel a « catalysé » les conflits épistémiques de son temps et en a dépassé les enjeux pour acquérir une autonomie pratique et théorique (Andreasen, 2007). Les critères, purement descriptifs et minimaux à l'origine (voir la conclusion de l'article de Feighner et collaborateurs de 1972 qui souligne justement le caractère « minimal » des critères proposés), ont été pris « trop au sérieux » par les cliniciens (Andreasen, 2007). Ce qui devait servir à favoriser la communication entre praticiens et de fondation stable dans le temps est devenu le principe canalisateur de la pratique psychiatrique. Certains auteurs ont avancé que le succès du DSM relevait également de son caractère « hégémonique » dans le paysage nosographique de la psychiatrie (Romelli et al., 2016). Ces auteurs soutenaient en effet que le DSM ne pouvait tolérer d'autres classifications. Cette affirmation

⁴² Ces précautions ne sont pas sans rappeler celles qui sont prises autour du projet RDoC, que nous étudierons par la suite.

est actuellement contestable. Nous reviendrons largement sur cette impossibilité de cohabitation avec d'autres classifications, qui semblait encore inimaginable il y a une décennie. Dans les chapitres à venir de ce travail, en soulignant toujours l'importance d'une approche par la catégorie, nous verrons que d'autres classifications ont su cependant trouver une place en psychiatrie – tout en admettant qu'une telle intégration est bien plus complexe qu'elle n'y paraît.

Dans le sens de cette « prise au sérieux » du DSM, à ce jour, les (plus jeunes) psychiatres et soignants utilisent les classifications qu'ils ont toujours connues depuis le début de leur pratique. Une part croissante de cliniciens n'ayant connu que le DSM pouvaient avoir, de principe, une certaine bienveillance à l'égard des catégories (Quang, 2015), qui seraient acceptées sans critique comme des entités « déjà-là », notamment par facilité, par exemple parce que le praticien se reposerait sur la classification pour administrer son traitement (Decker, 2013; Quang, 2015). Ces suppositions resteraient encore à démontrer, non pas en ce qui concerne l'absence de critique de la part des jeunes cliniciens (critique qui est indéniable, même si implicite), mais en ce qui concerne cette bienveillance « innée ». C'est sans doute pour une grande part le manque d'outils conceptuels qui empêche les cliniciens de se détacher des catégories du DSM. En allant plus loin encore et en s'attardant sur les conséquences inattendues du DSM, il apparaît que son accréditation implicite dépendait justement de sa puissance de régulation sociale. En effet, le DSM a permis de raffermir l'autorité sociale du soignant – qui détient à travers lui un savoir modulable. De ce fait, cette classification pourrait constituer un instrument scientifique de régulation, remis aux mains des praticiens. Le soignant aurait la capacité de juger de l'état psychique de son patient, avec toutes les conséquences qu'un tel jugement peut provoquer (obtention de compensations, aménagements, intégration dans des institutions, etc.).

Cependant, dès lors qu'on adopte une telle position apparemment dichotomique (entre un « bon » DSM et un DSM politisé « dévastateur »), il convient d'éviter certains écueils, que Schacht (1985) a su bien mettre en avant dans *DSM-III and the politics of truth* : ce n'est pas parce que le DSM est politisé que sa scientificité doit être remise en cause. Les variables de « scientificité » et de « politisation » sont en effet

potentiellement interdépendantes. La science ne peut s'extraire des contraintes extérieures qui la façonnent – et les implications sociopolitiques ne diminuent pas la valeur d'un outil constitué scientifiquement. La psychiatrie n'a jamais été isolée de la société et de la politique, tout comme il n'existe pas de science « pure » de l'influence du contexte social et tout comme les politiques publiques dépendent intimement des instances scientifiques pour rendre leurs décisions. La pathologisation des émotions aurait pu apparaître dans un contexte individualiste croissant, la consommation de psychotropes en lien avec les progrès techniques de la pharmacothérapie et la tendance réductionniste en rapport avec le développement institutionnel des centres de recherche en sciences fondamentales. Il ne s'agit pas tant de relativiser l'impact du DSM que de le replacer dans un environnement scientifique multifactoriel, dans lequel les questionnements de la psychiatrie étaient aiguillonnés par des faits sociaux et une forte évolution de la logique épistémologique en psychiatrie.

Conclusion

Le DSM est progressivement devenu un outil heuristique relativement fiable et offrant un langage « universel » pour les cliniciens. Il s'appuie sur une terminologie et un ensemble de concepts capables de répondre à des objectifs variés (scientifiques, économiques, administratifs et sociaux). Il a donc été également utilisé pour arbitrer des enjeux qui dépassent la recherche, comme celui de fournir à la psychiatrie un caractère plus scientifique. De plus, d'un point de vue de son histoire, le DSM a apporté de nouvelles valeurs à la psychiatrie, qui sont originellement celles de l'École de Saint-Louis et de la volonté d'opérationnaliser les catégories de troubles. D'un point de vue législatif et pénal, il a fourni un support de défense des patients, dans la mesure où la déresponsabilisation et l'expertise se fondent encore actuellement sur les critères du DSM. D'un point de vue économique, il a permis le remboursement et le

soutien des usagers de santé mentale, en fournissant un socle commun aux différentes instances administratives et financières chargées de ces tâches. D'un point de vue culturel et pragmatique, la désignation des maladies et l'apposition diagnostique peuvent également être considérées comme une plus-value : elles ont permis une redéfinition des maladies en santé mentale et ainsi une réappropriation de celle-ci par les patients (Demazeux, 2013a, p. 248). L'aspect athéorique du DSM (ou du moins sa volonté de se détacher de toute théorie) a pu jouer en sa faveur : il se présentait ainsi comme un document « œcuménique », capable d'accueillir non seulement les différentes orientations théoriques des cliniciens et chercheurs (on devrait plutôt parler de document « polythéorique »), mais également les disciplines adjacentes à la psychiatrie, telle que la sociologie, les théories de la communication, le béhaviorisme, la linguistique ou l'anthropologie. Les représentations sociales ont donc évolué en conséquence et avec elles les enjeux de stigmatisation, d'affiliation communautaire et de mobilisation individuelle qui se sont créés autour de cet outil. En étudiant le contexte historique de son développement, nous nous apercevons que le DSM peut répondre à de nombreuses fonctions qui dépassent largement les fonctions utiles à la pratique clinique. Nous avons compris dans ce chapitre que la volonté d'organisation standardisée et opérationnalisée des critères au sein de catégories constituait la clé de voûte de cette multiplicité fonctionnelle. Nous verrons dans la seconde partie de cette thèse que ces différentes fonctions cliniques et non-cliniques sont supportées principalement par l'importance donnée par le DSM à la notion de cause commune aux troubles psychiatriques, que nous discuterons plus largement sous le terme de « variable latente ». Construit sur le modèle médical tel que nous l'avons abordé au fil de ce chapitre, le DSM s'est en effet ancré sur la volonté de s'appuyer sur une telle variable latente – la catégorie diagnostique⁴³. Or,

⁴³ Le modèle médical conceptualise les symptômes comme les conséquences d'une entité pathologique sous-jacente (par exemple, l'humeur triste et l'anhédonie seraient les conséquences visibles d'une entité sous-jacente, à l'origine de la dysfonction – préjudiciable – qui constitue le trouble psychiatrique selon la HDA). Cette entité sous-jacente, qui n'est pas nécessairement physiologique mais peut-être statistique (comme la dimension d'une analyse factorielle) ou conceptuelle (Putnam, 1975 ; Wakefield, 1997, 1999), nous l'avons nommée variable latente, permettant de rendre compte de l'agrégation mutuelle des symptômes. Le modèle médical rend tous les symptômes à peu près également centraux au sein d'un trouble, car tous les symptômes résultent de l'affection sous-jacente. Autrement dit, sur le plan diagnostique, les symptômes deviennent théoriquement équivalents et interchangeables (Cramer et al., 2010; Lux & Kendler, 2010; Schmittmann et al., 2013).

comme nous le reverrons, la volonté de s'appuyer sur une telle approche catégorielle met en tension cette nosographie et sa potentielle pertinence clinique.

Pour conclure, le DSM a été construit et soutenu en tant qu'outil permettant de légitimer la psychiatrie dans le champ médical. Dans ce sens, il pourrait être considéré comme un relatif succès, au vu de son expansion dans le paysage médical contemporain. Mais une fois cette fonction de légitimation remplie, face à la complexité des tableaux cliniques que les cliniciens et scientifiques rencontrent dans leurs pratiques, persiste au moins une question fondamentale : en quoi le DSM peut-il être considéré comme un outil « utile » pour la pratique clinique, du moins utilisé isolément ? C'est au travers de l'étude d'au moins deux autres nosographies psychiatriques que nous pourrons indirectement répondre à une telle question. Nous verrons dans le chapitre suivant comment l'orientation qu'a emprunté le DSM dans son histoire pour tenter de se rapprocher d'une meilleure (et fiable) définition du trouble psychiatrique a conduit à l'émergence de critiques et à un virage vers la précision – exemplifié par le projet des *Research Domain Criteria*. Par la suite, dans la seconde partie de cette thèse, nous verrons que la définition plus ou moins explicite de ce qu'est un trouble psychiatrique pour le DSM (et ces autres nosographies que nous allons détailler) entre en confrontation avec une autre définition possible d'une espèce psychiatrique plus profondément ancrée dans la pratique clinique.

LE PROJET *RESEARCH DOMAIN CRITERIA* : DES NEUROSCIENCES À L'ENVIRONNEMENT

Le projet *Research Domain Criteria* (RDoC) constitue avant tout un programme de recherche permettant de structurer les expérimentations vouées à recueillir des données probantes issues de la neurobiologie. Mais au-delà de cette horizon pragmatique, s'incarnent au sein de sa « matrice » un ensemble de valeurs, méthodes et intérêts témoignant d'une conception bien particulière de la psychiatrie (et qui dépasse largement l'enjeu d'une « classification » proposée comme un « manuel ») – conception neurobiologique qui nous intéresse dans ce chapitre.

Le projet RDoC est loin de constituer un « changement de paradigme » ou une « révolution scientifique », selon les termes de ses promoteurs. Il constitue l'instanciation concrète d'une volonté de retour à la physio(patho)logie pour la psychiatrie. Avec les RDoC, la psychiatrie cherche à opérer un retour aux neurosciences, sans pour autant (naïvement) tenter de justifier les catégories diagnostiques (du DSM) par la physiopathologie. En pratique, les RDoC cherchent à mettre en relation des fonctions cérébrales ordonnées en domaines (ou construits et sous-construits, sur les lignes du tableau), à différentes échelles du vivant (allant du niveau moléculaire aux mesures auto-rapportées, dans les colonnes du tableau).

En ce sens, dans ce chapitre, nous entendons montrer que les RDoC constituent une approche à la fois pluraliste au sein des neurosciences, tout en étant réductionniste vis-à-vis de la psychiatrie clinique. Cependant, à rebours des conceptions faisant des RDoC une matrice absolument détachée de la psychiatrie clinique, nous montrons que le projet s'appuie et discute nécessairement avec le contexte socio-historique au sein duquel il se développe. Nous mettons ainsi en avant les enjeux d'applicabilité clinique et d'intégration temporelle et des déterminants environnementaux des RDoC. L'objectif de ce chapitre est de caractériser la mise en tension entre la volonté d'intégrer le projet RDoC à la pratique clinique psychiatrique et la nécessité de répondre aux multiples enjeux scientifiques et institutionnels qui ont été à l'origine même de ce projet.

Sur le plan philosophique et nosologique, nous montrons que le projet semble bien peu original par rapport à une conception de la psychiatrie comme celle du DSM. En effet, notre hypothèse est que les dimensions proposées par les RDoC, qui s'inscrivent dans une approche par la catégorie/dimension, soulèvent des problématiques au moins tout aussi difficiles à résoudre que les catégories du DSM. L'objectif critique de ce chapitre est de fournir un aperçu de ces dimensions/catégories au travers de l'identification de trois défis que tentent de relever le projet RDoC : la validité de construit de ses dimensions, son vœu d'intégrativité autour de la validité externe (voué à justifier l'existence des fonctions) et sa prise en considération de l'environnement. Nous analysons dans la seconde partie de cette thèse le projet RDoC en tant qu'approche fondée sur la dimension (plus que sur la « catégorie » à proprement parler, bien que les implications soient similaires), avec la volonté de fournir des éléments neurobiologiques justifiant l'existence de fonctions cognitives spécifiques. DSM comme RDoC cherchent à légitimer la psychiatrie en « opérationnalisant » (pour le DSM) ou en « fondant sur des preuves » (pour le projet RDoC) les troubles psychiatriques (pour le DSM) ou les « fonctions cérébrales » (pour le projet RDoC).

Introduction

En 2013, Thomas Insel (alors directeur du NIMH) et Bruce N. Cuthbert affirmaient que l'incapacité du DSM à soutenir la recherche et à développer des traitements efficaces contre les troubles psychiatriques justifiait un « changement de paradigme » dans la classification psychiatrique (Cuthbert & Insel, 2013). Ce changement s'est concrétisé sous la forme du projet RDoC, présenté dans la communauté scientifique à travers plusieurs articles dont celui de Cuthbert de 2014, publié dans *World Psychiatry*, qui s'intitule *The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology* – titre révélateur de ses ambitions⁴⁴. Dès 2009, il était présenté comme un « système de classification de la recherche » (ce qui, hors contexte, pourrait être traduit par un système de classification « pour la recherche », ou même « en recherche ») permettant d'*expliquer* les troubles mentaux selon les principales dimensions de la neurobiologie et du comportement observable (Cuthbert, 2014). Comme l'écrivaient Cuthbert et Insel en 2013, le projet RDoC fournit « des dimensions biocomportementales fondamentales qui se déploient à travers les catégories actuelles de troubles » (Cuthbert & Insel, 2013). Nous verrons par la suite la signification d'un tel déploiement « à travers » les catégories et ce que, dans ce contexte, représentent ces « catégories ». Le projet est initialement présenté comme une feuille de route cherchant à fournir un cadre dimensionnel pour mieux structurer la psychiatrie. Il repose sur l'identification de données probantes de différentes fonctions cérébrales, selon un système organisé en un certain nombre de niveaux d'explications.

⁴⁴ Le système est consultable dans son intégralité sur le site du NIMH <https://www.nimh.nih.gov/research-priorities/rdoc/index.shtml>. À noter qu'il existe d'autres cadres de ce type, comme la *NIAAA Addictions Neuroclinical Assessment* et la *NIDA Phenotyping Assessments Battery*, qui sont également des cadres diagnostiques dits « agnostiques » formulés par des groupes d'experts autour de systèmes biocomportementaux. De telles propositions ne sont cependant pas nouvelles dans l'histoire de la psychiatrie, comme nous le détaillerons plus bas.

Le projet RDoC est présenté par ses concepteurs comme la possibilité inachevée de structurer les bases de données de la recherche neuroscientifique (sur le rôle des bases de données, voir notamment *Integrative Neuroscience in Psychiatry: The Role of a Standardized Database* (Gordon, 2003)). Les RDoC souhaitent soutenir le développement des futures versions des systèmes de classification traditionnels que sont le DSM ou la CIM. Ils résonnent également avec un projet soutenu par l'Union Européenne, nommé ROAMER pour *Roadmap Mental Health Europe*, qui propose d'analyser les étapes nécessaires pour développer la recherche en neurosciences en Europe dans les prochaines décennies, cherchant à renseigner indirectement la clinique (Forsman et al., 2015). Certes, actuellement, les promoteurs des RDoC veillent à rester agnostiques quant aux types de diagnostics cliniques qui seront (éventuellement) en vigueur dans les décennies à venir et quant au potentiel d'applicabilité de cet outil. Ils « s'inscrivent dans la durée », en soutenant que l'agrégation de données valides constitue un projet constructif à *long terme* – potentiellement au prix d'un manque d'applicabilité à court terme. À la question de savoir comment « supporter » les décennies à attendre encore avant de voir apparaître un système constitué, Vilma Gabbay, un chercheur ayant reçu un financement de la NIMH pour un projet RDoC, déclarait : « Mais depuis combien de temps attendons-nous depuis Freud ? » (Bradley et al., 2017). En tant que « feuille de route neuroscientifique du NIMH » à destination de la psychiatrie, les RDoC pourraient être considérés au premier abord comme un simple programme de financement de la recherche. Nous soutenons cependant dans ce chapitre une justification épistémologique plus profonde du projet peut être menée, en cela qu'il représente l'instanciation concrète d'une certaine conception (pas uniquement contemporaine) de la psychiatrie. Autrement dit, même si leur portée pourrait être concrètement limitée à financer des projets de recherche, pour un nombre restreint de contributeurs

« élus » par le NIMH⁴⁵, l'ensemble des concepts et des valeurs qu'ils véhiculent s'ancrent dans une vision neurobiologique de la psychiatrie – et c'est cette vision qui nous intéresse ici. Notamment, nous nous intéressons au couplage de cette vision neurobiologique avec la nécessaire intégration clinique et des déterminants environnementaux du projet. Ce mariage « forcé » exemplifie à notre avis la nécessité pour toute conception de la psychiatrie de répondre à de multiples enjeux, qui n'a été pensée – du moins dans le cadre de ce projet – qu'en prenant appui sur la biologie.

Le fait que le projet RDoC ne soit pas voué à devenir à terme une *classification* est resté ambigu au fil des publications et des interventions de ses défenseurs (Clark et al., 2017). Dans les articles initiaux, notamment celui de Cuthbert (Cuthbert, 2014), l'auteur écrivait que le projet « pourrait devenir à terme » un système utilisable en clinique, laissant clairement entendre que cette utilisation pourrait devenir une classification à part entière. En ayant emprunté un virage computationnel avec son intégration au programme BRAIN 2025, le projet s'inscrit d'autant plus dans une rhétorique des promesses. Il est d'ailleurs légitime de se demander si l'ambition du projet d'embrasser la complexité des troubles psychiatriques sera à la hauteur de ses possibilités. Les résultats d'un « microscope mental » pour étudier la psychiatrie clinique, pour reprendre les mots de Posner et Raichle dans *Images of mind* (Posner & Raichle, 1994), seront-ils utilisables en pratique clinique, ou même utiles pour des recherches futures ? Les RDoC ne seraient-ils pas que des « *neurosciences in a dish* », pour reprendre la dénomination de l'article décrivant que l'essence de la schizophrénie avait été (enfin) fournie grâce à l'identification des

⁴⁵ Plus précisément, l'objectif pratique des RDoC est de fournir des financements pour favoriser la recherche autour des neurosciences. Cependant, la recherche utilisant comme point focal le cerveau n'est pas nouvelle : la leçon des décennies précédentes, montrant que malgré de nombreuses recherches et avancées en neurosciences, peu de découvertes ont eu de véritables implications en pratique clinique (Hall et al., 2017), ne semble pas avoir influencé les choix (neurocentrés) du NIMH. Un tel choix a des répercussions sur les autres domaines (par exemple les études de santé publique et portant sur l'environnement), puisque leur financement doit être réparti en fonction des programmes institutionnels prévisionnels. L'une des explications pouvant expliquer la persistance de ce choix neurocentré correspond à la notion de dépendance au chemin, c'est-à-dire à la persistance du développement dans une voie ou un champ disciplinaire pour lequel de nombreux investissements ont déjà été réalisés. Nous détaillerons le processus de la dépendance au chemin en annexe, montrant que l'intérêt pour les sciences du cerveau conduit nécessairement à négliger les interventions d'un autre genre (comme l'épidémiologie ou la psychologie sociale) (**Annexe 1 – « Dépendance au chemin et régime des promesses dans la nosologie psychiatrique »**).

cellules souches (« *Schizophrenia in a dish* » – (Callaway, 2011)) ? À ce propos, souvenons-nous de ce qu'écrivait Ramachandran il y a plus de vingt ans : « Les neurones miroirs feront pour la psychologie ce que l'ADN a fait pour la biologie : fournir un cadre unificateur et contribuer à expliquer tout un ensemble de capacités mentales qui sont restées à ce jour mystérieuses et inaccessibles à l'expérimentation » (Ramachandran, 2000). Ni les neurones miroirs (Forest, 2014; Southgate & Hamilton, 2008) ni la génétique en elle-même n'ont su convaincre les communautés de chercheurs qui travaillent dans ces deux domaines. L'une et l'autre de ces découvertes ont laissé un temps se confondre les hypothèses, les promesses, les croyances et les espoirs dans un cadre de travail unifié, mais finalement incapable de résumer la complexité du vivant. Ce qui était censé jouer le rôle de fondation des interactions sociales ou du déterminisme humain n'a pu répondre aux questions essentielles que se posait notamment la nosologie (Forest & Granger, 2015). À l'image des neurones miroirs ou de la génétique, les RDoC pourraient n'être qu'une nouvelle tentative de « résumer la complexité du vivant ». Certes, la matrice est plus grande, plus complète, plus exhaustive et plus interdisciplinaire que les modèles qui la précédaient. La démonstration de son utilité clinique semble loin d'être réalisée (Wakefield, 2014) et c'est ce que nous étudierons plus en détail dans ce chapitre. Dans de multiples interventions (consultables notamment sur le site du NIMH) et dans les articles postérieurs à 2015, comme celui de Cuthbert intitulé *The role of RDoC in future classification of mental disorders* (Cuthbert, 2020), le projet RDoC affirme n'être voué qu'à rester à l'état de construction dynamique en constante évolution, servant à fournir un cadre de travail intégrant des résultats neuroscientifiques. Quelques années après son lancement, l'ambition du projet est donc bien plus nuancée. Suivant l'état des connaissances et des publications les plus récentes, nous retiendrons dans un premier temps que le projet RDoC est avant tout destiné à la recherche. Il constituerait alors plutôt un ensemble de principes dynamiques permettant de « construire une base de connaissances cumulative sur la psychopathologie [émergeant] de perturbations dans le fonctionnement normal » (Lilienfeld & Treadway, 2016). Cependant, quelle que soit sa portée annoncée, nous considérons ce projet comme une conception de la psychiatrie parmi d'autres. L'évolution de ses

ambitions illustre combien cette conception est tiraillée entre son nécessaire socle neurobiologique et sa volonté de satisfaire les enjeux de pratique clinique. Nous montrerons l'importance de cette tension dans ce chapitre, en illustrant que les dimensions intégrées à la matrice tentent de rendre compte de la pratique clinique via différents procédés de validité, une tentative d'interprétation des unités d'analyse et une volonté d'ancrage dans une perspective environnementale.

Trois grands axes sont ainsi développés au sein des cinq sections de ce chapitre : i) l'importance de la notion de validité pour tenter de justifier la matrice des RDoC, en l'absence de tout ancrage clinique, ii) les enjeux d'intégrativité entre ses différents éléments et au sein de la communauté (neuro)scientifique iii) et enfin sa volonté d'intégrer la pratique clinique en considérant les comportements, le (neuro)développement et l'environnement (qui renvoie aux « déterminants environnementaux »). Nous verrons au travers de ces trois axes que c'est justement parce que les RDoC sont fondés sur une approche par la dimension qu'un certain constat d'échec devrait être tiré. En effet, en se structurant autour de fonctions psychologiques (appelées justement « dimensions » dans la matrice), les RDoC perpétuent l'approche par la catégorie du DSM : tous deux fournissent des « boîtes » (sous forme de construit psychologique pour les RDoC ou de catégories pour le DSM) dans lesquelles « ranger » des éléments (neurobiologiques pour les RDoC ou sémiologiques pour le DSM). Cette image de la « boîte » correspond justement à la notion de variable latente, qui renvoie à la définition d'une espèce psychiatrique en tant qu'espèce diagnostique. En cela, un certain nombre de difficultés de l'approche par la catégorie qu'amenait le DSM sont reconduites par les RDoC, en tant qu'approche par la dimension. Nous verrons dans la seconde partie de la thèse que ces approches par la catégorie / dimension sont relativement similaires (en termes d'inférences et d'enjeux) par rapport à ce que nous désignerons comme une approche par le symptôme.

Dans ce chapitre, plus précisément, nous détaillerons la matrice des RDoC dans une première section et nous verrons dans une deuxième section les facteurs ayant contribué au développement du projet RDoC. Dans une troisième section, nous explorerons la notion de validité de construit qui structure les RDoC.

En effet, nous choisissons de présenter la validité de construit avant toute autre chose car elle est centrale dans le projet, ayant été en partie à l'initiative de sa création et justifiant son existence et son expansion. Les RDoC promettent d'assurer la validité que la classification DSM n'aurait pas su apporter aux troubles psychiatriques. Comprendre les enjeux de la validité des RDoC est donc fondamental pour comprendre leur utilité, leur place dans la psychiatrie contemporaine et leurs implications pour les années à venir. Dans une quatrième section, nous verrons que la validité des RDoC ne pourrait être espérée sans que ne soit prise en compte l'intégration des unités d'analyse entre elles (les unités d'analyse correspondent aux colonnes de la matrice, qui représentent les différentes échelles du vivant, par exemple le niveau génétique ou le niveau moléculaire). L'intégration des unités d'analyses signifie la « prise en compte » d'un niveau (par exemple, cellulaire) par un autre niveau (par exemple, moléculaire). Cette prise en compte peut être (idéalement) comprise en termes ontologiques, de causalité, mais également (et plus simplement) en termes d'influence et d'association. Nous verrons que cette intégration est loin d'être triviale car les RDoC, au premier abord du moins, « agglomèrent » des unités d'analyse les unes à côté des autres sans présenter de franches relations entre ces unités. Nous verrons que cette intégration peut cependant se réaliser à d'autres niveaux de compréhension et notamment au niveau de l'interdisciplinarité (neuro)scientifique. Dans une cinquième et dernière section, en écho à cette tentative d'intégration des unités d'analyse entre elles, nous étudierons ce qui relève autant du « mythe » des RDoC que d'un long débat mené par certains opposants à ce projet. Il s'agira de comprendre pourquoi l'unité d'analyse « environnement » n'a pas été incluse à la suite des unités d'analyse portant sur les fonctions cérébrales (s'arrêtant aux « comportements auto-rapportés »).

I. Description de la matrice des RDoC

Il ne peut être nié que les RDoC se sont constitués en tant qu'initiative de financement de subventions de recherche, initiative dirigée vers un domaine très spécifique : les études des principales fonctions du cerveau. Clairement, il s'agit d'un projet constitué de promesses larges permettant de financer des recherches expérimentales étroites. Ainsi, il semblait important de présenter (initialement) le projet RDoC comme une « initiative de recherche » et de spécifier le contexte dans lequel il s'est déployé.

L'objectif initial des RDoC, en 2009, était de soutenir des programmes de recherches précis et de financer des études permettant d'intégrer, ensemble, les différents courants des neurosciences. Dans sa version la plus modeste, ses concepteurs espéraient que la littérature issue des recherches sur ce projet (et non le projet lui-même), portant sur les sciences du cerveau, pourrait être utilisée pour mieux expliquer les troubles psychiatriques (Cuthbert & Insel, 2010). Selon Cuthbert, les RDoC constituent donc un programme dynamique qui continue à évoluer au fur et à mesure de l'accumulation de nouvelles données de recherche (Cuthbert, 2014). De manière plus ou moins explicite et avouée, les bénéfices en semblaient alors relativement lointains : la « révolution nosologique » (Tabb, 2020) ne serait que le fruit de ces recherches et pourrait n'en être qu'un bénéfice secondaire. Sur le site du NIMH, on peut retrouver la description de la constitution des ateliers d'experts. Plusieurs années ont été nécessaires pour constituer et animer cinq ateliers constitués chacun de 30 à 40 experts⁴⁶ et correspondant à chacun des domaines qui forment la base des RDoC. Lors de ces ateliers, les participants (psychiatres et chercheurs en santé mentale pour la plupart) ont défini exactement quelles étaient les principales dimensions cérébrales (ou « fonctions », ou « domaines ») étudiées et validées dans la littérature (c'est-à-dire pour lesquelles des données probantes ont été rapportées).

⁴⁶ Leurs tâches étaient de déterminer quelles dimensions devaient être incluses dans chaque domaine, fournir une définition pour chaque dimension et proposer une liste des éléments utiles à chaque dimension, qui pourraient servir à la mesurer, en s'appuyant sur plusieurs unités d'analyse (Cuthbert & Insel, 2013).

La matrice des RDoC consiste en un tableau, ou « matrice », organisant les disciplines des neurosciences. Les lignes de la matrice sont des construits (ou constructions) renvoyant aux principales fonctions comportementales. À la suite de la description riche et détaillée donnée dès 1955 par (Cronbach & Meehl, 1955), l'article de Sanislow et collaborateurs de 2010, intitulé *Developing constructs for psychopathology research: Research Domain Criteria*, explique l'utilisation de ce terme de « construit », terme issu de la psychologie et des statistiques et qui peut s'appliquer à tout aspect spécifique du fonctionnement psychologique (Sanislow et al., 2010). Il s'agit de la « meilleure conception » d'une fonction (les fonctions sont listées dans la matrice – voir ci-dessous), bien que cette meilleure conception ne puisse jamais correspondre parfaitement aux mécanismes sous-jacents. Les construits n'ont pas pour vocation de s'accorder avec les catégories du DSM. Ils n'incluent volontairement pas la sémiologie traditionnelle (les signes et les symptômes), le terme « symptôme » n'étant (quasiment) jamais employé par les concepteurs ou chercheurs utilisant les RDoC. L'idée d'un tel cadre de travail est de faire en sorte que l'éparpillement des données neuroscientifiques soit régulé par l'existence de la matrice. Pour chacune des fonctions identifiées par les experts de chaque atelier, des données neuroscientifiques empiriques seraient rangées par « unité d'analyse ». Chaque unité d'analyse correspond à une colonne de la matrice (lorsque chaque fonction correspond à une ligne de la matrice). Autrement dit, le projet RDoC est voué à étudier les troubles psychiatriques en fonction des dimensions du comportement observable (les fonctions, en lignes) et des mesures neurobiologiques (les unités d'analyse, en colonnes). Cinq grands domaines de fonctionnement cérébral et comportemental sont identifiés dans les lignes de la matrice, auxquels pourraient s'ajouter d'autres domaines au fil du temps (par exemple, un sixième domaine lié au système sensori-moteur a été ajouté par la suite) :

- Les systèmes de valence positifs (comme la récompense et les habitudes) ;
- Les systèmes de valence négatifs (comme la réaction aux menaces, aux souffrances, aux situations aversives ou négatives) ;

- Les systèmes de valence cognitifs (comme la mémoire, l'attention ou les fonctions exécutives) ;
- Les systèmes des processus sociaux (comme la reconnaissance des émotions ou la théorie de l'esprit) ;
- Les systèmes d'éveil / de modulation (comme les rythmes biologiques et les systèmes d'éveil)⁴⁷.

Les colonnes de la matrice, qui correspondent aux « unités d'analyse » à différentes échelles du vivant (des gènes aux mesures auto-rapportées), reflètent les différents niveaux d'organisation de la recherche sur les domaines du fonctionnement cérébral et comportemental. Comme nous l'avons vu, l'approche scientifique du projet RDoC concorde avec l'héritage intellectuel de la neuropsychiatrie classique. Nous pouvons retrouver dans le contenu de ces fonctions l'héritage historique que nous avons décrit dans la section I de ce chapitre. Par exemple, le modèle de Meynert-Wernicke pourrait être retrouvé dans l'unité d'analyse nommée « Circuits », concrétisant la recherche historique d'un substrat dysfonctionnel localisé pour la psychopathologie. De même, des zones cérébrales comme les noyaux basolatéraux de l'amygdale sont décrites en écho aux travaux de Kleist (ou reflétant son influence, suivant le contexte dans lequel on étudie l'ancrage des RDoC) (Neumärker & Bartsch, 2003). Une colonne supplémentaire, intitulée « paradigmes », décrit les procédures expérimentales permettant de produire, détecter et mesurer les fonctions comportementales correspondant aux constructions de la matrice. Les « paradigmes » de la matrice des RDoC, situés au sein de la colonne la plus à droite, correspondent aux « tests expérimentaux » neuroscientifiques. Il peut s'agir par exemple d'un test d'orientation spatiale pour évaluer les capacités visuo-perceptives ou d'un test de motricité pour évaluer les capacités d'agentivité.

⁴⁷ Notons que les RDoC n'incluent pas de représentations détaillées des phénotypes cliniques, comme l'agressivité, le narcissisme, la labilité émotionnelle, les groupes sociaux, le vieillissement, etc., qui sont courants en recherche et en clinique mais qui ont peu (ou pas) « d'accent » biologique. Un cadre de travail complémentaire ou d'autres domaines pourraient être nécessaire pour les intégrer (Kirmayer et al., 2015).

Le terme de paradigme provient du vocabulaire de la recherche expérimentale : chaque domaine neuroscientifique utilise une série de tests expérimentaux spécifiques. Par exemple, la génétique utilise plutôt des paradigmes d'hybridation *in situ* ou des analyses biomoléculaires tandis que la neuroimagerie utilise plutôt des paradigmes testant les fonctions cognitives, comme le « *Go No Go* » ou le « *Stroop test* », deux tests permettant d'évaluer l'inhibition. La volonté des RDoC, en plus de rassembler au sein d'une matrice unique toutes les fonctions à toutes les échelles, est également de collecter ces tests expérimentaux issus de différentes traditions neuroscientifiques.

Il a été admis par les promoteurs des RDoC que les unités d'analyse ont été incluses car les résultats qui peuvent y être intégrés sont finis et que leur dimensionnalité est connue. À l'inverse, l'environnement reste potentiellement « infini » (Kendler et al., 2020). L'inclusion est donc évaluée en fonction de la faisabilité, selon un consensus scientifique et selon les réponses possibles découlant de la recherche. Nous soutenons ici que les domaines (ou construits, ou sous-construits, ou fonctions, ou dimensions, en lignes) n'ont été développés que pour répondre à un impératif pratique d'opérationnalisation des unités d'analyse (les échelles du vivant, en colonnes). En effet, les « fonctions » ne sont que des labels dont la consistance n'est apportée que par les données rangées au sein de la matrice. Ils ne possèdent qu'une valeur heuristique pour aider les différentes communautés neuroscientifiques impliquées dans la recherche en psychiatrie à ordonner les données. En tant que labels, ces constructions ne sont donc pas vouées à être opérationnalisables en clinique. C'est tout l'enjeu de l'argumentation de ce chapitre qui cherche à montrer qu'en tant qu'approche par la dimension, le projet RDoC présente des fonctions sans profondeur conceptuelle, puisqu'elles ne servent qu'à ordonner des données. La **Figure 2** présente les différents éléments de la matrice des RDoC.

		UNITÉS D'ANALYSE						
DOMAINES / CONSTRUITS	Gènes	Molécules	Cellules	Circuits	Physiologie	Comportement	Mesures auto-rapportées	Paradigmes
Système de valence négative								
Menace aiguë (« peur »)								
Menace potentielle (« anxiété »)								
Menace soutenue								
Perte								
Frustration non récompensée								
Système de valence positive								
Motivation d'approche								
Réactivité initiale à la récompense								
Système d'apprentissage								
Habitude								
Systèmes cognitifs								
Attention								
Perception								
Mémoire de travail								
Mémoire déclarative								
Comportement de langage								
Contrôle cognitif								
Système des processus sociaux								
Affiliation / attachement								
Communication sociale								
Perception / compréhension de soi								
Perception / compréhension des autres								
Éveil / Systèmes modulateurs								
Éveil								
Rythmes biologiques								
Veille-sommeil								

Figure 2 – La matrice des *Research Domain Criteria* (RDoC). En lignes, les principales fonctions cérébrales ; en colonnes, les échelles du vivant (« unités d'analyse »). La dernière colonne sert à ordonner les tests expérimentaux utilisés pour remplir la matrice (« paradigmes »).

II. Genèse des RDoC

1. L'influence du contexte scientifique de la première moitié du XX^e siècle

Derrière le projet de rassembler au sein d'une matrice différentes échelles du vivant, on retrouve dans le projet RDoC un programme cherchant à fonder la psychiatrie sur un ensemble de bases neurobiologiques et neuroanatomiques. Le projet RDoC s'inscrit dans la continuité d'un programme neuroscientifique débuté au moins au XIX^e siècle. Ce programme pourrait être distingué en deux mouvements. Le premier cherche à fonder la psychiatrie sur un ensemble d'éléments neuroscientifiques. Le second, plus tardif et apparu après 1980, est celui d'une tentative de prise de distance avec la réification des catégories du DSM. Ces deux mouvements ne sont pas indépendants : vouloir appuyer la psychiatrie sur des fondations cérébrales tient à une volonté de distanciation avec la réification opérée par le DSM. Zandersen et Parnas ont conclu leur article de 2019 par ce constat (après avoir démontré que la grande majorité des patients porteur d'un trouble de personnalité de type état-limite répondaient aux critères d'un trouble du spectre de la schizophrénie) : « Depuis le DSM-III, les diagnostics psychiatriques se sont réifiés et ont été considérés comme des espèces naturelles... Il faudrait peut-être plutôt réinstancier la psychopathologie théorique et empirique au cœur de la psychiatrie scientifique » (Zandersen & Parnas, 2019). Un projet de philosophie comme celui de Jonathan Tsou est justement de « ré-insuffler » de la biologie dans les catégories cliniques. Les troubles au sein du DSM sont parfois traités comme s'il s'agissait de véritables entités existant indépendamment d'un évaluateur (Hyman, 2010). Nous verrons ainsi par la suite que

l'abandon de la réification des troubles et la volonté d'ancrer scientifiquement la psychiatrie sur des étiologies ou sur des marqueurs cérébraux sont intimement liés.

Les RDoC sont donc le fruit d'une longue tradition (neuro)scientifique. Dans une discussion en ligne publiée sur le site internet des RDoC⁴⁸, Cuthbert met en avant combien « notre système diagnostique [qu'est le DSM] dépend réellement de notre héritage culturel occidental ». Par-là, il signifie que la médecine occidentale s'est construite sur une association d'effets et de causes et plus proche de nous, sur le modèle médical (que nous détaillerons largement par la suite). Une telle médecine, construite sur des effets et des causes, rend compte de manifestations symptomatiques témoignant de l'altération de mécanismes physiologiques sous-jacents (c'est-à-dire renvoyant à une ou plusieurs cause(s) sous-jacente(s)). En nous reposant sur cette conception de la médecine, nous détaillerons dans cette partie l'ancrage historique qui a permis la constitution des RDoC.

À partir du XIX^e siècle, la psychiatrie a été identifiée comme une spécialité médicale autonome (Kendler, 2009). Pour des cliniciens comme Philippe Pinel, la nosologie doit être fondée sur la sémiologie, c'est-à-dire sur les symptômes et les regroupements de symptômes (Telles-Correia & Sampaio, 2016). C'est dans la seconde moitié du XIX^e siècle que, soutenue par l'essor de recherches scientifiques capables d'isoler des « causes » neuroanatomiques et neurobiologiques (par exemple, voir l'article de 1970 de Robins et Guze, de l'École de Saint Louis, dans lequel les auteurs soutiennent que la validité de la catégorie des schizophrénies nécessite une explication causale) (Robins & Guze, 1970), les maladies psychiatriques ont été progressivement intégrées au sein des sciences du cerveau (Forest, 2014). Ainsi, pour des auteurs comme Wilhelm Griesinger, Carl Wernicke et Alois Alzheimer « les maladies mentales sont des maladies du cerveau » (ici, selon les termes de Griesinger) (Arens, 1996). En prenant l'exemple de Carl Wernicke, neurologue célèbre pour ses travaux sur les aphasies (Forest, 2015), un modèle étiopathogénique général de la maladie mentale a été proposé sur la base de la correspondance d'une manifestation pathologique (l'aphasie) et d'une lésion localisée dans le cerveau. Il émit l'hypothèse qu'il

⁴⁸ Consultable sur : <https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/index.shtml>.

pourrait en être de même pour les maladies mentales. En s'appuyant sur les travaux de Theodore Meynert (Deacon, 1989) portant sur les schémas de connexion des symptômes neurologiques et psychiatriques, Wernicke a ainsi créé un « modèle local fonctionnel » et proposé une « hypothèse de la jonction », selon lequel les symptômes psychopathologiques proviendraient de l'interruption des connexions associatives au sein du cerveau (Whitaker & Etlinger, 1993). En 1934, issu du même genre de recherche cherchant à « localiser la psychopathologie », Karl Kleist, un élève de Wernicke, a conçu une carte de localisation cérébrale selon l'hypothèse du « localisationnisme » (Guelfi, 2010), ancrant les fonctions mentales dans le cerveau, à l'image de ce qu'ont fait quelques années auparavant, en 1909, les cartes neuroanatomiques de Brodmann et en 1925 celles de von Economo (von Economo & Koskinas, 1925). Selon certains auteurs, le localisationnisme de Kleist et de von Economo partagerait de nombreuses similitudes avec les résultats obtenus en imagerie par résonance magnétique fonctionnelle et en tomographie par émission de positons, comme l'ont souligné Neumarker et Bartsch dans *Karl Kleist (1879–1960) – A Pioneer of Neuropsychiatry* (Neumärker & Bartsch, 2003) (**Figure 3**).

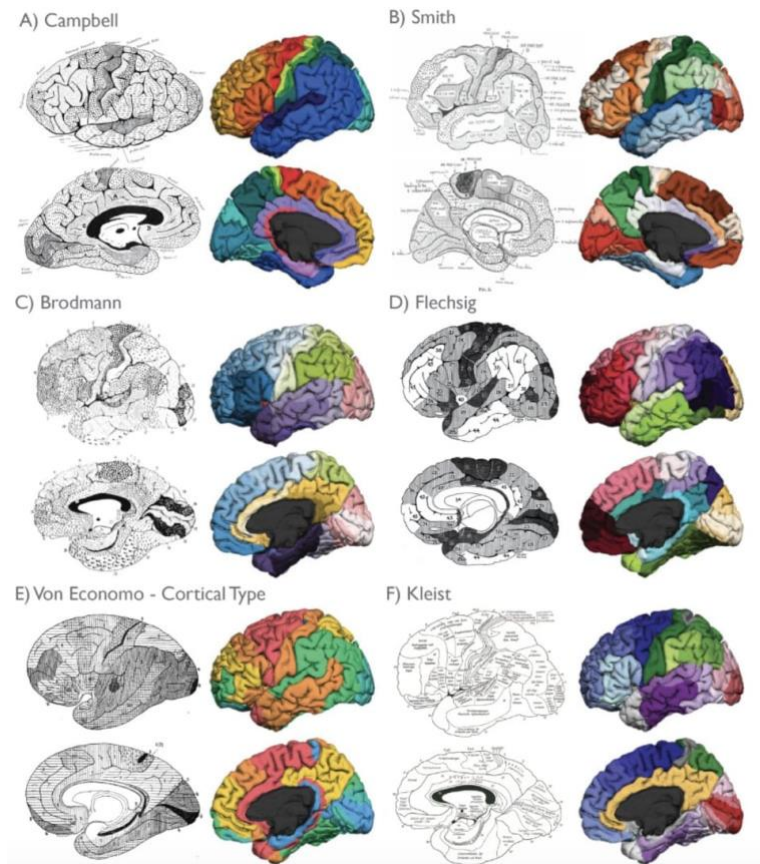


Figure 3 – Localisation des fonctions corticales selon l'architecture cérébrale. De A. à F., les représentations issues des différents atlas de Campbell, Smith, Brodmann, Flechsig, Von Economo et Kleist (Pijnenburg et al., 2021).

Ces modèles tendent à localiser les troubles psychiatriques au sein du cerveau. Ils reflètent l'orientation d'une psychiatrie en transition vers la neurologie, de la fin du XIX^e siècle jusqu'au milieu du XX^e siècle. Cependant, ces modèles se sont heurtés à la difficulté pratique de déceler une (unique) cause neuroanatomique, à l'image de ce qui a pu être fait initialement, par exemple, lors de la mise en évidence d'agents infectieux, dans la maladie d'Alzheimer (Hippius & Neundörfer, 2003) ou dans la toxoplasmose ou la syphilis (Kendler, 2005). L'incapacité de ces modèles neuroanatomiques et neurobiologiques à fournir des résultats cliniques aurait pu contribuer à éloigner le débat de l'explication causale et

mécaniste en psychiatrie vers la sémiologie, la phénoménologie ou la psychanalyse (Ellenberger, 1981; Jaspers, 1963; Shepherd, 1982). Nous verrons ici, en nous appuyant sur l'exemple du projet des RDoC, que cet éloignement n'a pas du tout eu lieu.

Pour cela, nous développerons brièvement trois modes de pensée psychiatriques de la première moitié du XX^e siècle qui ont cherché à harmoniser leur objet d'étude et leur méthode d'investigation pour améliorer l'unité de la psychiatrie. Nous verrons dans les sections suivantes que c'est bien le premier mode de pensée, axé autour de la neuropsychiatrie, qui a pris le pas sur les autres dans les précédentes décennies. Le premier mode de pensée, neuropsychiatrique, proposait pour objet d'étude le cerveau. De fait, il utilisait des méthodes d'investigation qui devaient être adaptées à cet organe (Forest, 2014). Le deuxième mode de pensée, en prenant l'exemple des travaux menés par Karl Jaspers en phénoménologie (Jaspers, 1963), fait de la psychiatrie un objet des sciences humaines, dont la méthode d'investigation devrait passer par une collecte de signes et de symptômes intégrés à un système philosophique (Ellenberger, 1981; Tsou, 2011). En prenant l'exemple des psychiatres cliniciens de la première moitié du XX^e siècle comme Emil Kraepelin, Kurt Schneider et Eugen Bleuler, le troisième mode de pensée serait critériologique et concevrait une psychiatrie fondée sur la description des diagnostics (Demazeux, 2019). Nous détaillons rapidement cette troisième approche de la psychiatrie du début du XX^e siècle en prenant l'exemple de la schizophrénie. Les catégories de schizophrénie et de trouble bipolaire proviennent du travail de psychiatres allemands comme Eugen Bleuler et Emil Kraepelin, Kraepelin ayant proposé de distinguer les troubles psychotiques des troubles bipolaires sur la base d'une distinction liée à leur origine cérébrale. Ainsi, la schizophrénie et les troubles bipolaires ont été distingués l'un de l'autre sur l'hypothèse de causes sous-jacentes différenciées. Ce bref exemple montre que la standardisation des classifications psychiatriques traditionnelles comme le DSM (qui intègre donc la schizophrénie et les troubles bipolaires) est partiellement fondée sur l'hypothèse que des causes sous-jacentes, distinctes, sous-tendent les différentes catégories de troubles (Klerman, 1977; Murphy, 2013). Et comme nous le verrons, il existe de nombreux arguments permettant d'affirmer que la conception de

la psychiatrie critériologique (celle du troisième mode de pensée, comme on la retrouve dans le DSM) peut difficilement être séparée d'une conception de la psychiatrie fondée sur les causes (celle du premier mode de pensée) (voir notamment la contribution de (Tsou, 2015), *DSM-5 and psychiatry's second revolution: Descriptive vs. theoretical approaches to psychiatric classification*).

2. Le contexte scientifique de la seconde moitié du XX^e siècle : une triple déception

À partir de la seconde moitié du XX^e siècle, le développement des neurosciences et des sciences du comportement a permis de rapprocher la psychiatrie des sciences naturelles et en particulier des sciences du cerveau (Forest, 2014). Ainsi, une discipline hybride comme la psychiatrie, à l'intersection des comportements sociaux, du langage, des interactions de groupe et en même temps des sciences du cerveau, nécessiterait d'utiliser une approche hybride entre sciences humaines et sciences naturelles (une telle approche hybride a été soutenue par exemple par (Weber, 1904)). Cette période a donc marqué la volonté de réunir les différentes disciplines qui gravitent autour de la psychiatrie (comme la neuropsychiatrie, la phénoménologie ou la psychiatrie descriptive). Cependant, nous verrons que ce programme unificateur n'est en aucun cas un projet abouti. Les sciences du cerveau ont pris la plus large part de ce projet. Nous développerons par la suite quels sont les principaux facteurs ayant contribué aux progrès des neurosciences en psychiatrie. Nous détaillerons trois influences majeures des RDoC qui ont conduit à proposer un tel système conceptuel cherchant à intégrer les troubles psychiatriques à des dimensions neuroscientifiques : celle provenant d'une psychiatrie biomédicale déçue par les catégories traditionnelles du DSM (influence nosologique), celle provenant de la difficile application des catégories traditionnelles du DSM à la clinique (influence clinique) et celle provenant des déceptions de la recherche elle-même (influence de la recherche).

Premièrement, la déception liée à la nosologie psychiatrique traditionnelle aurait été symboliquement à l'origine de la proposition des RDoC. La volonté de révision nosologique du DSM semblait nécessaire au début des années 2000, principalement pour des doutes concernant la pertinence des catégories. En effet, ces dernières ne fournissent pas d'entités (ni marqueurs) suffisamment précises pour guider la recherche et les approches thérapeutiques en pratique clinique. Les catégories du DSM ne correspondent pas aux cibles moléculaires, génétiques et morphologiques du cerveau et les traitements correspondant à ces catégories sont ubiquitaires et non spécifiques. Compte tenu des progrès des neurosciences et des sciences du comportement, entre 2002 et 2013, le DSM-5 était attendu comme un manuel fortement influencé par l'approche des sciences du cerveau (Hyman, 2007). Des séries de conférences ont été réalisées et des groupes de travail ont été constitués afin d'intégrer les nouvelles données issues des neurosciences (Ayearst et al., 2012; Swedo et al., 2012). C'est en tout cas ainsi que fut mené le début du projet, même si nous resterons prudents sur ces allégations. Au cours de cette période pendant laquelle fut constitué le DSM-5 (de 2002 à 2013), plusieurs laboratoires du NIMH ont évalué des données issues de la neuroimagerie, de la génomique, de la biologie moléculaire et des sciences du comportement : les résultats de ces études ne s'alignaient pas sur les diagnostics catégoriels du DSM (Kozak & Cuthbert, 2016). Des perturbations dans les mêmes fonctions cérébrales cognitives, émotionnelles, perceptives ou comportementales (par exemple, des systèmes associés à la peur, à la récompense ou aux processus cognitifs) pouvaient être atteints dans plusieurs troubles (par exemple, l'état de stress post-traumatique ou même le trouble dépressif caractérisé), alors que tous les patients d'une catégorie diagnostique donnée ne présentaient généralement pas les mêmes perturbations dans un même système. En d'autres termes, les catégories DSM représentaient des syndromes hétérogènes par rapport aux fonctions cérébrales cognitives, émotionnelles, perceptives ou comportementales. Ces difficultés avaient bien été anticipé par un certain nombre d'auteurs (par exemple, Dominic Murphy dans son ouvrage *Psychiatry in the Scientific Image* en 2006 (Murphy, 2006), ou dans un de ses articles de l'*Oxford textbook of philosophy and psychiatry* de 2006 également). Mais en même temps, les progrès d'une « psychiatrie de précision »

ont influencé le développement des RDoC. Celle-ci peut être définie comme un ensemble de développements scientifiques contemporains destinés à améliorer la compréhension et le soin clinique à partir d'une amélioration de la prise en compte des spécificités individuelles (Giroux & Gauld, 2023). C'est en partant de ce constat sur l'hétérogénéité du DSM que le NIMH a lancé un appel pour élaborer un plan stratégique, en 2008, susceptible d'étudier les troubles mentaux en fonction des dimensions (ou fonctions) du comportement observable et des mesures neurobiologiques. Cet appel a abouti à l'initiative des RDoC (Cuthbert & Insel, 2010). Comme nous l'avons vu, ces ambitions, malgré quelques raffinements par rapport au DSM-IV-TR, ont été revues à la baisse : le DSM-5 s'est révélé conservateur et soumis à la dépendance de son chemin passé (Demazeux & Singy, 2015). C'est entre autres cette déception d'une psychiatrie qui se voulait biomédicale et qui n'a pu aboutir en pratique, qui a conduit dès le début des années 2000 à imaginer un autre système de conception des troubles capable d'intégrer les progrès des sciences du cerveau (Cuthbert, 2014).

Deuxièmement, la difficile application des catégories traditionnelles du DSM à la clinique a également contribué à cette déception et à ce virage vers une nosographie neuroscientifique. En effet, dès la parution du DSM-III, les cliniciens ont pu se montrer « frustrés » du manque de performance clinique du DSM (Cuthbert, 2014). Les cliniciens ne s'appuient pas directement sur les catégories du DSM, mais ils ont tendance à utiliser le « même ensemble restreint de catégories diagnostiques » (Maj, 2020) comme moyen d'obtenir des remboursements pour les patients, ainsi que pour communiquer avec d'autres cliniciens – en utilisant notamment la bonne fiabilité de ses catégories (voire notamment *Classification for clinical practice: How to make ICD and DSM better able to serve clinicians* de (First & Westen, 2007)). Bien souvent, en clinique, il semble que les diagnostics ne sont pas posés parce qu'ils correspondent aux phénomènes cliniques perçus précisément par les cliniciens, mais bien parce qu'ils facilitent leurs décisions de soin, elles-mêmes fondées (avant tout) sur leur expertise individuelle (Bowker & Star, 1999). En acceptant implicitement cette conception « déflationniste » vis-à-vis de la classification (Bolton, 2012), les RDoC balaient les catégories traditionnelles pour tenter de proposer un

système fournissant une correspondance entre des marqueurs biologiques⁴⁹ (à plusieurs échelles d'analyse) et des comportements observés (dont témoignent les principales fonctions).

Troisièmement, en plus de l'influence liée aux déceptions émanant de la nosographie traditionnelle et à l'importance de l'influence de la clinique par rapport à la faiblesse des catégories, l'essor des RDoC pourrait provenir (apparemment paradoxalement) des faiblesses et limites de la recherche en psychiatrie et en neurosciences. Au début des années 2000, la psychiatrie d'orientation biomédicale affirmait que la révision du DSM-IV-TR devrait se fonder sur les sciences du cerveau. Cette affirmation semble bien peu originale au vu de l'histoire de la psychiatrie. Pourtant, elle paraît tout aussi ambitieuse que lorsque Robins et Guze, dans leur article de 1970 (intitulé *Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia*), affirmaient que « l'apparente schizophrénie de bon pronostic était une autre maladie » et que cette distinction devait être faite grâce aux travaux « génétiques, psychodynamiques, cliniques, sociologiques, chimiques, physiologiques ou thérapeutiques » (Robins & Guze, 1970). Autant de domaines cités dans les années 1970 qui rappellent, curieusement, ceux que nous détaillerons par la suite en présentant la matrice des RDoC, développée presque un demi-siècle plus tard.

Comme l'écrivaient Wong et collaborateurs en 2010, « En faisant correspondre les bons traitements aux bons patients, on pourrait passer d'une thérapie universelle à une thérapie ciblée et augmenter le rapport bénéfice-risque pour les patients » (the LHA genetics group et al., 2015), remarque sous-tendant le développement d'une psychiatrie de précision et d'un remaniement des protocoles d'études en recherche. Si la psychiatrie se tourne vers cette « médecine de précision », la recherche, s'appuyant (et fondée) nécessairement sur les catégories du DSM, ne satisfait pas pleinement les acteurs de la psychiatrie (Tabb, 2020). Ce constat a par exemple été clairement visible du fait de l'inadéquation des catégories avec la génétique psychiatrique. Ainsi, les déceptions rencontrées par la neurogénétique psychiatrique au début

⁴⁹ Nous verrons plus loin comment ces « marqueurs » doivent être caractérisés : sont-ils des signatures, des facteurs étiologiques, des mécanismes ?

des années 2000 ont déplacé l'intérêt du NIMH vers les neurosciences morphologiques et fonctionnelles (c'est-à-dire décrivant les structures cérébrales et leur dynamique au sein du cerveau) et en particulier vers le connectome⁵⁰. Le connectome correspond à la carte des connexions neuronales dans le cerveau (son « schéma de câblage ») (Sporns et al., 2005). Au début des années 2000, les technologies d'étude de la morphologie fonctionnelle et dynamique du cerveau étaient en effet en pleine expansion et promettaient alors un nouveau « modèle local fonctionnel » selon l'ambition de Wernicke (1874) (Wernicke, 1874). Cela pourrait d'ailleurs être une des raisons pour laquelle le niveau de la connectivité cérébrale constitue bien une des échelles du vivant intégrée au RDoC, alors que l'échelle « génétique » en a été retirée. Finalement, la promesse de précision et le désir de faire « revivre » l'approche biomédicale aurait pu soutenir le déploiement du projet RDoC.

Nous allons voir que la genèse et le développement des RDoC sont aussi liés à des progrès scientifiques apparus dans la psychiatrie contemporaine des dernières décennies.

3. L'influence de la technologie et des marqueurs biologiques

L'émergence des RDoC ne dépend pas que de déceptions nosologiques, cliniques ou issues de la recherche. Nous fournirons deux exemples paradigmatiques qui aident à comprendre l'origine de leur émergence. Premièrement, les technologies ont éclairé d'une autre manière les faiblesses du DSM ; deuxièmement, l'hétérogénéité d'une nouvelle proposition comme celle des RDoC a inmanquablement conduit à reconsidérer ses catégories. Depuis les années 2000, la psychiatrie contemporaine est entrée dans une ère où les techniques de neuroimagerie et de neurogénétique permettent de confirmer les

⁵⁰ En parallèle de ce déplacement scientifique, on observe un déplacement sociologique : il existe un réaligement des intérêts des chercheurs et au fur et à mesure des échecs des promesses (Rüppel, 2019). En effet, les chercheurs font de nouvelles promesses au fil des déconvenues, répondant ainsi à la nécessaire dynamique scientifique qui se construit sur ses déceptions et insuccès.

prédictions et les hypothèses faites lors des siècles précédents concernant le fonctionnement du cerveau. La validité des catégories traditionnelles (qui n'avait pas été directement testée lors de la constitution du DSM, contrairement à la fiabilité, avec les études sur le terrain, ou « *field trials* » – (Clarke et al., 2013)) a été remise en question au fur et à mesure, notamment, des progrès technoscientifiques (Zachar et al., 2014). Ce problème pourrait être posé différemment en affirmant que l'espoir promis par les technologies et les expérimentations a été déçu. L'absence d'un test diagnostique ou d'une thérapie qui s'appuierait à la fois sur la neuroimagerie et sur les catégories du DSM est le témoin actuel de ce manque de validité. Nous pourrions avancer deux hypothèses concernant cette inadéquation entre catégorie et marqueur : soit ce sont les marqueurs apportés par les technologies qui ne correspondent pas aux catégories, soit ce sont les catégories qui ne sont pas valides.

Concernant la première hypothèse, si les catégories cliniques correspondaient parfaitement à la génétique ou aux données de l'imagerie, c'est-à-dire si elles avaient une bonne validité (notamment de construit – (Cronbach & Meehl, 1955)), des tests et des traitements basés sur la génétique ou l'imagerie auraient pu être théoriquement développés. Du moins, il s'agit d'une hypothèse optimiste fondée sur l'observation d'une augmentation croissante des progrès scientifiques et technologiques des dernières décennies.

La deuxième hypothèse concernant le manque de validité des catégories diagnostiques pourrait paraître triviale : en effet, les concepteurs du DSM(-III) n'ont pas directement cherché à valider les catégories. Ou du moins, ils ont *indirectement* voulu les valider à l'aide de validateurs externes (c'est-à-dire d'éléments extérieurs à la classification, comme la génétique ou les études familiales). C'est cependant cette absence de validité qui a conduit à élaborer un nouveau programme donnant la priorité aux technologies d'imagerie et à la biologie.

Pour illustrer cette difficulté à identifier une relation marqueur-catégorie, prenons l'exemple des traitements pharmacologiques. Ceux-ci fonctionnent de manière ubiquitaire. Cela signifie qu'ils agissent sur un nombre important de troubles distincts sans être spécifique à un unique trouble. Ce caractère ubiquitaire va dans le sens d'un décalage entre les catégories du DSM et les substrats cérébraux sous-

jacents. En 2010, des scientifiques du laboratoire pharmaceutique d'AstraZeneca ont écrit un article intitulé « *The Drug Hunter's Perspective* » qui montre qu'en moyenne « un médicament psychiatrique commercialisé est efficace dans environ la moitié des cas » – et un grand nombre d'études dans la dépression admet que l'efficacité pharmacologique ne dépasse pas 30 % de l'intensité de la symptomatologie (Wong et al., 2010). S'il est (historiquement) délicat de dire que la psychopharmacologie en est « à ses débuts⁵¹ », il paraît également vain d'espérer une concordance fine entre un marqueur biologique et une unique catégorie du DSM. En fait, nous admettons que ces deux hypothèses (concernant la validité des catégories diagnostiques et la relation entre un marqueur et une catégorie) sont toutes deux valables. Ces discordances ne proviennent pas uniquement d'imperfections technologiques (hypothèse 1) ou d'imprécisions nosographiques (hypothèse 2). Elles proviennent également de la nature même des troubles psychiatriques, nature complexe qui explique pourquoi il faut considérer en même temps ces deux hypothèses. Afin d'analyser la nature de ce qui est désigné par trouble psychiatrique, il nous faut nécessairement faire un retour sur la notion de « trouble » et sur celle de « catégorie » telle qu'elle est décrite dans le DSM, comme vu dans le chapitre précédent. Plus précisément, les frontières d'un trouble, tel que défini par le DSM, pourraient être considérées comme floues (Haslam, 2002). Un trouble donné ne présente pas de correspondance avec un marqueur donné. Donnons en un exemple. Dans la revue *Nature*, en 2011, le premier rapport sur les cellules souches utilisables dans le spectre de la schizophrénie a été publié sous le nom de *Schizophrenia in a dish*. L'enthousiasme de la communauté scientifique a été lié non pas tant au fait qu'une nouvelle thérapeutique était proposée dans le cadre de la schizophrénie, mais que, pour la première fois, une thérapie moléculaire génétique allait pouvoir être « proposée en clinique » (Callaway, 2011). Nous sommes aujourd'hui parfaitement conscients de la naïveté concernant les espoirs de l'application directe d'une thérapie (moléculaire, génétique ou même morphologique, c'est-à-dire basée sur l'imagerie

⁵¹ La chlorpromazine, le premier traitement médicamenteux de la psychiatrie « contemporaine » synthétisé le 11 décembre 1950, a été commercialisé sous le nom de « Largactil » pour souligner l'aspect « large » (utilisé dans la racine *Larg-*) de son spectre d'action.

cérébrale) à la pratique clinique. Un certain nombre d'auteurs, tels que Michael Owen, un psychiatre généticien affilié à l'Université de Cardiff, exprimaient la diversité des manifestations des troubles schizophréniques : « Ces troubles ne sont pas vraiment des troubles. La schizophrénie n'existe pas. C'est un syndrome. C'est un ensemble de choses que les psychiatres ont regroupées ». Jeffrey Poland proposait également en 2006 de « dépasser le concept de schizophrénie », soulignant de la même manière que la schizophrénie ne pourrait être qu'une « idée reçue » sans crédibilité scientifique.

Nous pourrions détailler brièvement ces termes de « syndrome » et de « trouble », car ils concernent indirectement les enjeux des RDoC (Hall et al., 2015; Owen et al., 2011). Pour certains, un trouble renvoie à une dysfonction (« mentale ») qui entraîne des perturbations dans le comportement du sujet (voir par exemple le traité de *Psychiatrie* de 1926 de (Codet, 1926, p. 12), ou (Wakefield, 1992)). Une catégorie du DSM, pour laquelle n'existe aucune « essence » sous-jacente au niveau cérébral (ou même à d'autres niveaux explicatifs) est avant tout un syndrome, c'est-à-dire un ensemble de phénomènes (par exemple, des symptômes) que les psychiatres ont regroupés. À ce syndrome s'ajoute dans la catégorie des indicateurs de fréquence, de rythme, de durée, d'exclusion d'autres catégories, etc. Alors pourquoi nomme-t-on une catégorie du DSM un « trouble » ? C'est là toute l'ambiguïté du manuel. Certes, une réponse triviale serait que l'ajout des spécificateurs (de durée, de fréquence, d'exclusion, d'âge, etc.) permettrait de différencier le syndrome (ensemble de symptômes) du trouble (ensemble de symptômes et des spécificateurs). Cependant, c'est sur cette fine distinction que tout la logique (« athéorique ») du DSM s'appuie. Le trouble psychiatrique décrit dans le DSM désigne un syndrome pour lequel on espère trouver une cause, notamment parce que les manifestations de ce trouble sont réunies ensemble parce qu'une dysfonction (conçue selon la psychologie évolutionniste, pour Wakefield) permettrait d'en expliquer la présence commune (au sein de construits) (nous détaillerons précisément ces points dans un chapitre dédié aux symptômes) (Wakefield, 1992; Wakefield & First, 2013). Si l'étiologie (et non la

cause, comme nous le verrons) était découverte, le trouble porterait l'intitulé de « maladie »⁵². Mais en l'absence de notification de telles étiologies, les catégories du DSM semblent vouées à rester des syndromes au sein de catégories, dont la définition conceptuelle (comme celle proposée par Wakefield) permet de parler de « troubles »⁵³. C'est ainsi que dans leur version initiale (au moins non computationnelle⁵⁴) les RDoC contournent la question des syndromes. Ils proposent plutôt d'associer au marqueur une fonction comportementale, dégagée de toute catégorie. Cuthbert, en 2017, expliquait dans une conférence : « Essayons de regarder ce que le cerveau fait réellement et quels en sont les principaux systèmes, afin d'essayer de penser aux troubles mentaux en ces termes, au lieu de construire des

⁵² Nous parlons ici de maladie au sens médical du terme, dans la mesure où il s'oppose conceptuellement au trouble. Il est important de ne pas confondre cette terminologie avec celle de la maladie entendue comme *sickness* par rapport à la pathologie entendue comme *disease* (Boorse, 1977).

⁵³ Cependant, à quoi bon parler de trouble ou de syndrome si ni l'un ni l'autre ne permet de guider l'action en pratique ? Par exemple, aucune thérapie de type « solution miracle » (« *magic bullet* » en anglais) n'est applicable pour ses troubles-syndromes, comme le souligne Healy, en expliquant que le désir de trouver de telles solutions est au cœur même de la recherche sur les psychotropes (Healy, 1998). Il n'est pas anodin de remarquer que cette expression (« *magic bullet* ») est également utilisée en oncologie, concernant la possibilité d'associer une tumeur à un théranostic, c'est-à-dire un marqueur associé à un traitement ciblé – car dans cette discipline, de telles « solutions miracles » ont été trouvées (quoique dans de rares cas ciblés).

⁵⁴ Dans le sens d'une convergence entre RDoC et psychiatrie computationnelle, plusieurs initiatives ont vu le jour. Dans la branche interventionnelle de la division du NIMH portant sur la recherche translationnelle (*Division of Translational Research*), un programme de psychiatrie computationnelle (*Computational Psychiatry Program*) s'appuie explicitement sur le projet RDoC : celui-ci est présenté comme soutenant « la recherche translationnelle utilisant des modèles computationnels pour valider les construits des RDoC en clinique ». Il cherche à combiner des outils mathématiques et computationnels avec des techniques neurophysiologiques, neuroanatomiques, neurochimiques et / ou de neuroimagerie. L'objectif principal du travail de validation des RDoC par la psychiatrie computationnelle passe par la comparaison des construits présents au sein de la matrice. Il s'agit donc d'une mise en relation et comparaison des unités d'analyse et construits – comparaison qui n'avait pas été réalisée dans la proposition initiale de la matrice. Cette recherche d'intégration correspond à une volonté de comprendre les troubles psychiatriques en tant que systèmes complexes, c'est-à-dire d'entités en interaction, connectant des éléments multiples, interdépendants et dynamiques. Trois questions sont explicitement posées dans ces propositions de financement : « Les différents domaines de comportement sont-ils séparés les uns des autres ? », « Dans quelle mesure reposent-ils sur des circuits neuronaux distincts ou se chevauchant ? » et « Quelles sont les relations entre les domaines, les construits et les sous-construits, à la fois en termes de structure corrélacionnelle et de circuits neuronaux sous-jacents ? ». Il s'agit de transformer la matrice des RDoC en une matrice « relationnelle et hiérarchique » : celle-ci permettra de relier différentes informations entre unités d'analyse et entre niveaux, en analysant les relations complexes (au sein et) entre les circuits impliqués dans les différentes fonctions. Il existe également une initiative émanant du plan stratégique du NIMH (*National Institute of Mental Health Strategic Plan for Research*), paru en 2020. Celui-ci intègre une série d'initiatives précédentes qui proposaient, de manière plus ou moins centrale, de considérer les neurosciences et la psychiatrie computationnelle et notamment : le *Grand Challenges in Global Mental Health Initiative*, le *Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative strategic plan* (avec le *BRAIN 2025: A Scientific Vision*, le *BRAIN Initiative 2.0: From Cells to Circuits, Toward Cures* et le *BRAIN Initiative and Neuroethics: Enabling and Enhancing Neuroscience Advances for Society*). Il est intéressant de remarquer que Joshua Gordon, le directeur du NIMH, s'est d'ailleurs très tôt intéressé aux approches computationnelles et a rédigé en 2016 un ouvrage intitulé *Computational Psychiatry* (Redish & Gordon, 2016).

catégories de symptômes en supposant que les problèmes cérébraux et les problèmes de comportement puissent correspondre à ces catégories définies en l'absence totale de connaissances sur le cerveau il y a 100 à 2 000 ans⁵⁵ ».

Pour conclure cette section, il semble qu'une des principales raisons des limites diagnostiques et thérapeutiques de la psychiatrie contemporaine corresponde au regroupement artificiel de phénomènes hétérogènes (symptômes / comportements) en des troubles. Mais en même temps, dans l'histoire des RDoC, c'est justement la présence de ces regroupements qui, pour une grande part, a contribué à la volonté de développer un cadre de travail tel que le projet RDoC.

4. Vers un nouveau système

« Les RDoC tentent de raffiner le dilemme existentiel auquel est confronté la psychiatrie : celle-ci pourrait être l'un des domaines médicaux les plus convaincants si les troubles mentaux étaient compris comme des troubles cérébraux, mais les troubles cérébraux semblent relever de la neurologie »

(Wakefield, 2014)

Comme le décrivaient Cuthbert et Insel en 2013 dans leur article intitulé *Toward the future of psychiatric diagnosis: the seven pillars of RDoC*, sept éléments distinguent la démarche des RDoC de celle des nosographies traditionnelles comme le DSM (Cuthbert & Insel, 2013). Ces éléments sont tantôt des réalités, intrinsèques aux RDoC (comme la dimensionnalité, l'aspect protocolaire, les preuves scientifiques des construits), tantôt des promesses non encore réalisées (volonté de recherche translationnelle, mesures fiables et valides de la psychopathologie, poids équivalents entre dimensions du vivant ou relations entre RDoC et DSM). En résumant ces éléments, nous verrons que le projet RDoC

⁵⁵ Disponible en ligne sur : <https://www.youtube.com/watch?v=xFmpdKagDy4>.

tente de se défaire d'un certain nombre de limites reliées aux catégories du DSM, en retombant finalement sur des « dimensions » qui amènent des problèmes plus ou moins similaires.

Les RDoC proposent en effet une perspective de recherche translationnelle, qui tend à dépasser les catégories du DSM (via les construits de la matrice) tout en ayant une utilité clinique. La recherche translationnelle se fonde sur le fait que les sciences fondamentales, par exemple la génétique, les neurosciences et les sciences du comportement, pourraient servir de point de départ pour prendre en considération les tableaux cliniques. Cependant, les RDoC proposent de nouveau de valider ces construits au prisme de la clinique, puisque malgré les données neurobiologiques, l'aspect translationnel ne pourrait être rempli que si ses construits sont pertinents cliniquement. De plus, ils proposent une approche dimensionnelle de la psychopathologie. L'aspect dimensionnel du projet ne s'arrête pas à l'évaluation de la gravité mais cherche à inclure toute la gamme de variations comportementales allant du normal au pathologique, prenant ainsi en compte les deux extrêmes d'une dimension. Par exemple, selon certains modèles neuroscientifiques, la manie serait reliée au système de récompense (puisque celui-ci serait dérégulé dans le cadre de la manie : l'individu ayant une augmentation de la recherche du plaisir et de la récompense, conduisant au phénotype comportemental maniaque). Ainsi, la manie serait l'extrémité « haute » de la dimension de recherche de récompense et l'apathie représenterait l'extrémité « basse » de la dimension de recherche de récompense. Il s'agit donc d'essayer de « faire rentrer » les troubles psychiatriques « dans la boîte » d'une psychologie comportant plusieurs niveaux de problèmes : mais une telle approche peut-elle espérer être éclairante pour la psychiatrie ? Et est-ce qu'un tel découpage en dimensions est valide ? Ces points sont d'autant plus intéressants que ce vœu de dimensionnalité constitue un échec partiel dans l'élaboration du DSM-5 et un succès théorique dans l'élaboration d'une classification que nous étudierons par la suite, la HiTOP. Cependant, la question du contenu des dimensions des RDoC se pose comme cela a été le cas dans le cadre du DSM. Ensuite, le projet RDoC cherche à développer des mesures fiables et valides. L'optimisation des marqueurs permettraient une meilleure compréhension des troubles, une meilleure prévention et des traitements plus

ajustés, notamment en améliorant le seuil, ou « point de bascule », du normal au pathologique. Les RDoC proposent aussi un protocole général pour les études. Les protocoles se font en deux étapes. Ils débutent par un « cadre d'échantillonnage » [*sampling frame*], qui évalue quels patients peuvent être inclus dans une étude en fonction de ou des dimension(s) prises en considération. Il utilise un groupe témoin pour valider le fait que la gamme étudiée soit « complète ». Dans la deuxième étape, l'étude prend la forme d'une étude traditionnelle dans laquelle les cohortes fondées sur les catégories traditionnelles sont remplacées par les dimensions choisies. Ce point semble bien plus central dans l'entreprise des RDoC qu'il ne le laisse paraître. En effet, les RDoC ne sont pas uniquement un simple programme de financement, mais bien la concrétisation d'une conception de la psychiatrie relativement ancienne. De plus, le projet offre une structure intégrative accordant un poids égal aux différentes échelles d'explication, tant au niveau des unités d'analyses à plusieurs échelles du vivant qu'au niveau du raffinement des fonctions comportementales. Nous discuterons de cette spécificité « intégrative » par la suite. Rajoutons qu'il n'intègre que les construits pour lesquels il existe des preuves solides fondées sur des données probantes. Au départ du moins, l'objectif des RDoC était de « fournir une base solide sur laquelle acquérir de l'expérience », sans chercher à inclure toutes les pathologies répertoriées dans les différentes catégories du DSM. Au fil de ses développements cependant, la matrice a de plus en plus cherché à explorer ses propres recouvrements avec une grande part de la psychiatrie clinique (et donc avec les différentes catégories). Il existe ainsi un écart bien réel et difficilement surmontable, entre les données probantes existantes en recherche (intégrées explicitement aux RDoC) et les nécessaires données qui devraient être recueillies en pratique clinique (non intégrées explicitement aux RDoC). Enfin, le projet RDoC cherche volontairement à ne pas relier les dimensions à des définitions fixes des troubles. En cela, il semble que le projet RDoC ne modifie pas substantiellement le système de santé mentale fondé sur les catégories traditionnelles. On retrouve ici une des ambiguïtés majeures du projet RDoC, qui propose à la fois une « révolution nosologique » et finalement se positionne comme un système adjacent, parallèle au DSM et qui tente de ne pas « ébrécher » ses catégories. Encore une fois,

on retrouve dans cette ambiguïté la volonté de faire table rase, tout en ne touchant pas aux objets déjà existants de la psychiatrie. L'ensemble de ces éléments tempère les ambitions des RDoC et laissent craindre que cette proposition fondée sur des dimensions ne constitue qu'un palliatif biologique du DSM et de ses catégories. Mais voyons désormais en pratique ce que contient le projet.

Nous proposons d'illustrer la structure de la matrice des RDoC par une métaphore. Le cerveau serait comme une région traversée par des routes de différentes tailles. Chacune remplirait une fonction fondamentale en permettant de relier deux zones fonctionnelles en termes de comportements. Il existe ainsi un ensemble de circuits principaux pour répondre aux situations menaçantes, un autre ensemble de circuits pour réagir à un stress ou pour obtenir de la nourriture, pour se souvenir ou planifier ses actions et des circuits qui permettent de réguler la veille et le sommeil. Globalement, en conservant cette métaphore de la circulation routière, s'il y a un embouteillage sur l'autoroute parvenant à Paris, il se peut que le problème soit un accident survenu beaucoup plus bas, en lien avec, imaginons, de graves problèmes de circulation à la sortie de Lyon. Mais il peut également s'agir d'un accident à l'entrée de Paris. Déceler l'origine de l'embouteillage correspondrait à identifier les circuits cérébraux qui sont à l'origine de l'anomalie. Cette métaphore pointe un fait essentiel : il n'existe pas d'axe qualitativement « pathologique » en soi, mais seulement des avaries plus ou moins graves (« proximales ») du cerveau. Cette constatation justifierait le développement d'un programme de recherche s'attachant à déterminer la gamme complète de variations cérébrales du normal au pathologique (Lilienfeld & Treadway, 2016). Pour le DSM, la méthode de validation des troubles psychiatriques a été historiquement inversée par rapport à un tel système. Premièrement, il s'agit de comprendre un trouble psychiatrique de manière syndromique, en observant, recueillant et interprétant les signes et les symptômes du patient pour les inscrire dans des catégories. Ce n'est que secondairement et sur la base de ces constatations sémiologiques que les dysfonctionnements sous-jacents sont recherchés. Pour une partie des disciplines scientifiques adjacentes à la psychiatrie, cette recherche s'intègre aux sciences du cerveau. Dans la

conception des RDoC, fondée sur des « voies » neuroscientifiques, ce sont les gammes de fonctionnement comportemental qui sont considérées dans un premier temps. Ces fonctions évoluent sur un continuum du normal au pathologique et sont pour cela nommées des « dimensions » (Helzer et al., 2009). Les comportements seraient expliqués par de telles dimensions, permettant ainsi de concevoir les troubles mentaux comme des ensembles de comportement dont le caractère pathologique dépendrait d'un seuil défini sur un continuum et, fait important, ces comportements seraient désignés par des notions et un vocabulaire issu uniquement des neurosciences. Expliquons mieux pourquoi ce genre de conception des troubles mentaux cherche à ancrer les fonctions cérébrales le long d'un continuum. Ce système aurait-il pu créer des construits catégoriels au sein desquels chaque fonction cérébrale serait ou non atteinte (selon la dichotomie dysfonctionnel / non dysfonctionnel) ? Nous nous appuyerons sur un exemple pour répondre à cette interrogation. Nous pourrions imaginer un segment horizontal, sur lequel les cas « historiquement typiques » du trouble bipolaire (sur la base de la nature des symptômes) seraient disposés à l'extrémité droite et les cas « historiquement typiques » de la schizophrénie à l'extrémité gauche. Emil Kraepelin, utilisant cette dichotomie, dirait que chaque patient devrait être soit porteur d'une bipolarité, soit d'une schizophrénie, même si les quelques patients situés au milieu du segment pourraient être catégorisés comme « schizoaffectifs ». Quel que soit le diagnostic, chaque patient ne pourrait recevoir qu'un unique diagnostic dont la distinction dépendrait de sa position sur le segment. Ce modèle ne peut permettre de concevoir les troubles mentaux et leur complexité. En analysant les données empiriques, on doit reconnaître qu'il existe une dispersion des patients tout au long de ce segment de droite. Certains sont plutôt situés à l'extrémité gauche, possédant des symptômes de bipolarité, mais cela n'empêche en rien qu'ils comportent également des symptômes de schizophrénie⁵⁶. Dichotomiser de la sorte les troubles psychiatriques est difficile à assumer pour le clinicien qui se confronte à la réalité clinique. Par exemple, concernant le choix du traitement, une telle distinction est

⁵⁶ Nous verrons dans le chapitre sur la HiTOP que l'aspect dimensionnel peut également être détaillé au niveau des symptômes eux-mêmes, qui évoluent le long d'un spectre de gravité. En effet, ici, nous parlons de dimension en désignant des troubles, des catégories ou des fonctions, rassemblant des ensembles de symptômes.

rarement réalisée par le clinicien. Devant un tel cas « historiquement atypique », celui-ci se soustrait aux catégories traditionnelles pour traiter directement les symptômes et les plaintes de la personne. Il ne lui importe pas plus de connaître la « cause » du trouble, considéré comme le reflet d'une anomalie naturelle sous-jacente, que de fournir une étiquette pratique à un ensemble d'éléments⁵⁷. En cela, la psychiatrie est une des disciplines dont l'approche symptomatique peut être cliniquement plus pertinente que l'approche étiologique (notamment parce que les prises en charge sont plus adaptés aux symptômes décrits qu'à une étiologie supposée, comme nous le reverrons dans la seconde partie de ce travail). Pour les patients situés au centre du segment de droite, (qui ne possèdent pas plus de symptômes schizophréniques que de bipolarité), aucune préconisation ne peut être faite, posant tout à la fois des problèmes de santé publique (par exemple, « Que rembourser comme pathologie au long cours ? ») que des questions éthiques (par exemple, « Comment annoncer un diagnostic ? ») ou expérientielles, concernant l'incertitude diagnostique du patient lui-même (par exemple, « Pourquoi suis-je différent ? »). Ainsi, même un patient avec un trouble psychiatrique « pur » au sens catégoriel et historique du terme pourrait avoir un fonctionnement qui ne nécessiterait aucune prise en charge. De plus, une telle dichotomie entre catégories diagnostiques ne permet pas de prendre en compte les différences interindividuelles de fonctionnement quotidien. En effet, des signes et des symptômes exprimés très intensément chez un patient peuvent ne pas être préjudiciables ; des symptômes peuvent être exprimés moins intensément chez un autre patient mais être très handicapants (préjudiciables) en termes de fonctionnement quotidien. Ce point a fait l'objet d'une littérature croissante en philosophie de la psychiatrie, interrogeant la question du seuil et du potentiel préjudice « porté » par les symptômes (Cooper, 2005, 2013; Wakefield & First, 2013). Nous reviendrons dans la seconde partie de cette thèse sur cette notion de préjudice, de souffrance et d'acceptabilité intégrées à chaque symptôme.

⁵⁷ Nous insisterons sur ce point dans le chapitre sur les symptômes, en montrant que le psychiatre recueille avant tout un préjudice vécu et exprimé par le patient.

Les RDoC cherchent à relever le défi de la dimensionnalité. En proposant des dimensions à la place des catégories, ils cherchent à concevoir, appréhender et prendre en charge les troubles psychiatriques, en se rapportant avant tout à un fonctionnement cérébral indépendant des catégories du DSM. Cependant, si l'entité de base des RDoC est la dimension, qui nécessite qu'un certain seuil de trouble soit délimité, une question fondamentale se pose : qu'en est-il du seuil du normal et du pathologique ? Pour répondre à ce genre de question sur le normal et le pathologique au sein des RDoC, nous pourrions nous appuyer sur les développements de Jerome Wakefield, qui affirmait que les RDoC constituent une tentative de « raffiner le dilemme existentiel auquel est confrontée la psychiatrie » en faisant des dysfonctionnements cérébraux les principaux domaines et substrats de la psychiatrie (Wakefield, 2014). Wakefield poursuivait en citant Wittgenstein : « En psychologie, il existe des méthodes expérimentales et une confusion conceptuelle... L'existence d'une telle méthode expérimentale nous fait penser que nous avons les moyens de résoudre les problèmes qui nous préoccupent ». Wakefield semblait présenter le projet RDoC comme l'expression paradigmatique des préoccupations de Wittgenstein : la crainte d'un ambitieux programme de recherche empirique (s'appuyant sur les « méthodes expérimentales ») mais dont le cadre conceptuel serait dangereusement fragile. Wakefield affirmait en effet que le projet RDoC n'aurait aucun moyen de distinguer des limites des troubles psychiatriques (il parle alors de validité conceptuelle, renvoyant à la distinction entre trouble et absence de trouble) (Wakefield, 2007; Wakefield & First, 2013). Nous soutenons entièrement cette proposition : les RDoC ne permettent pas de distinguer le normal du pathologique, ce n'est ni leur volonté, ni dans leur capacité. Ce point illustre un fait doublement important : non seulement les RDoC sont fondés sur une approche dimensionnelle, qui fait se rejouer toutes les questions de pertinence clinique des catégories du DSM, mais ces dimensions perdent en plus leur capacité de discrimination du normal et du pathologique.

Cette nouvelle difficulté est assumée par le projet RDoC, qui tente de se décaler de cette question. La dimensionnalité est une propriété fondamentale des RDoC. Elle se retrouve à la fois dans l'indice de gravité de la dysfonction (la dysfonction est conçue selon un continuum de gravité) et dans la gamme de

fonctionnement de la population générale (autrement dit, la continuité est également défendue au niveau de la santé de la population). Une grande partie de la psychopathologie pourrait être placée le long d'un continuum. De ce fait, les tenants des RDoC déclarent « éviter d'avoir à définir les troubles », pour plutôt se concentrer sur les données capables d'améliorer « les révisions des futures versions du DSM et de la CIM en incluant des seuils quantitatifs et empiriques pour les troubles » (Drury & Cuthbert, 2015). Nous retrouvons là explicitement le vœu de ne pas séparer le normal du pathologique. Cependant, une autre version de considération pour le normal et le pathologique au sein des RDoC pourrait être donnée. Comme nous l'avons souligné précédemment, c'est en raison des progrès technologiques récents en neurosciences et en neurogénétiq ue que la volonté de fonder les troubles psychiatriques sur des dysfonctions cérébrales s'est posée de manière si aiguë. Les RDoC avancent en effet qu'en comprenant mieux ce continuum entre fonction et dysfonction, via des processus expérimentaux, il sera(it) alors possible d'améliorer la différenciation entre ce qui est un trouble de ce qui n'en est pas un. Mais la question se pose de savoir si la possibilité d'utiliser des « méthodes expérimentales » peut suffire à édifier un cadre de travail permettant de différencier les troubles psychiatriques des conditions non pathologiques. D'ailleurs, historiquement, l'existence d'un programme de recherche empirique fondé sur la méthode expérimentale semble avoir été une revendication aussi ancienne que la revendication de délimiter le normal du pathologique. Claude Bernard, déjà, se retrouvait face à la question de la dimensionnalité décrite entre le normal et le pathologique pour chaque fonction physiologique. Que pourrait apporter de plus un projet comme celui des RDoC ? Nous tenterons de répondre à ce genre de questions en détaillant la matrice des RDoC.

III. Validité de construit des RDoC

Nous avons vu que les RDoC sont construits sur la base de dimensions, sous forme de fonctions (rendant compte de « dimensions fonctionnelles »), qui permettent l'évaluation de la psycho(patho)logie le long d'un continuum. Cependant, ces dimensions amènent au moins trois défis que tentent de relever les RDoC, que nous allons détailler : 1) la validité de construit ; 2) l'intégrativité ; 3) et la prise en compte de l'environnement dans la matrice lorsque de telles fonctions sont considérées. Nous analyserons dans cette section la notion de validité telle qu'elle est entendue au sein du projet RDoC, celle de « construit » et, de fait, la validité de construit. Dans la suite de la thèse, nous reviendrons à plusieurs endroits sur cette notion de validité, tantôt pour en renforcer la portée, tantôt en fournir une approche critique.

1. Validité

Les RDoC se structurent autour d'une matrice intégrant les résultats des neurosciences au moyen de différents paradigmes expérimentaux. Cette matrice constitue un cadre de travail présentant des construits supposés « valides ». Nous proposons d'employer ici une définition originale et relativement modérée de la validité, en tant qu'elle constitue un processus de mesure de concordance entre deux éléments (classiquement entre un outil est « ce qu'il est censé mesurer », mais plus concrètement entre deux outils ou, comme nous le verrons plus loin, entre des éléments non directement mesurables⁵⁸).

De leur côté, les catégories du DSM sont censées répondre à des objectifs cliniques, par exemple diagnostiques ou thérapeutiques. Ces catégories ne peuvent être cependant considérées comme « réalistes » (voir notamment les travaux de Sullivan – (Kincaid & Sullivan, 2014; Mattu & Sullivan,

⁵⁸ Cette définition peut même être étendue au temps : la validité correspond à la concordance entre cet outil à un temps donné et ce même outil à un temps plus éloigné, montrant ainsi sa pertinence par la persistance de l'utilisation de l'outil par la communauté scientifique et / ou clinique. Par extension, ce que nous appelons « concordance » peut également être compris par contraste, dans le cadre d'un type de validité dite « divergente » : est valide un outil qui n'est pas concordant avec un outil qui n'est pas censé mesurer ce qu'il mesure.

2021; Sullivan, 2017a, 2017b)), car elles ne cherchent pas à établir une « correspondance directe » entre leurs critères et la « réalité » des phénomènes cliniques, pour reprendre les termes utilisés par (Hempel, 1965), témoignant ici d'une volonté de validité. Fiabilité et validité sont donc mise en tension par une « contrainte pragmatique⁵⁹ » (que la classification DSM remplit en proposant des critères permettant d'améliorer la prise de décision scientifique et clinique) et de la « contrainte réaliste » (témoignant d'une volonté de rapprocher une catégorie d'une description clinique ou scientifique et qui nécessite un appui de la classification sur des données probantes) (Kincaid & Sullivan, 2014). Les RDoC cherchent à concilier ces deux contraintes. D'un côté, ils répondent à la contrainte pragmatique en proposant un cadre de travail homogène constitué de données empiriques, qui se voudrait fiable. De l'autre, ils cherchent à répondre à la contrainte réaliste en faisant en sorte que ce cadre de travail et les construits qu'il contient soient « valides » – encore une fois, ici, au sens trivial d'un rapprochement entre le construit et la réalité conçue comme le résultat du croisement de plusieurs perspectives et communautés de recherche. Mais qu'est-ce donc que cette validité que tente de proposer les RDoC ?

La notion de validité est essentielle pour comprendre le projet RDoC. Celui-ci constitue une réorientation de la recherche qui détourne l'intérêt des « clusters de symptômes cliniques s'appuyant sur le consensus » (Cuthbert & Insel, 2013) vers une recherche qui concevrait la psychopathologie comme impliquant une « combinaison complexe de perturbations dans des processus plus fondamentaux ou des dimensions fonctionnelles ne correspondant pas nécessairement aux catégories de troubles actuellement identifiées » (Kerns et al., 2004). Cuthbert écrivait en 2014 que pour être inclus dans la matrice des RDoC, « les construits dimensionnels devaient être *validés* en tant que fonctions comportementales : il devrait exister la preuve de l'existence d'un circuit ou d'un système neuronal jouant un rôle prépondérant dans la mise en œuvre de cette fonction » (nous soulignons) (Cuthbert, 2014). Le construit dimensionnel n'est validé en tant que « fonction » que si des données (neuroscientifiques) sont jugées satisfaisantes.

⁵⁹ Le terme « pragmatique » est entendu au sens large, dans la mesure où il aide à la prise de décision clinique et scientifique.

La matrice des RDoC cherche donc avant tout à mettre en avant des construits validés au sein de plusieurs unités d'analyse et grâce à différents paradigmes expérimentaux. Les propositions d'étude menées dans le cadre du projet RDoC augmentent la possibilité de valider les construits qui le structurent. Elles ont des critères d'inclusion plus larges que celles utilisant des catégories psychiatriques classiques, les variations dans le fonctionnement de systèmes neurocomportementaux peuvent être étudiées chez les patients comme chez les personnes « sous-symptomatiques » (prise en compte systématisée en recherche d'une population qui ne pouvait l'être avec les classifications traditionnelles et des relations entre les niveaux d'analyse peuvent être proposées (par exemple en mettant en avant les interactions des gènes avec des molécules dans le cadre de la recherche sur des traitements). Du fait de cette volontaire souplesse, les RDoC cherchent à répondre à différentes interrogations méthodologiques, nosologiques et épistémologiques (continuum normal et pathologique, fonctions cérébrales multiples, aspect transdiagnostique par rapport aux catégories du DSM, relations entre niveaux d'analyse, etc.).

2. Validité de construit

Le projet RDoC met en jeu des construits dimensionnels. Les deux notions de validité et de construit sont reliées historiquement au sein du concept psychométrique de « validité de construit ». Nous ne tenterons pas dans cette section de résumer la notion isolée de « construit », qui a fait l'objet d'une intense littérature depuis (au moins) (Cronbach & Meehl, 1955). Nous discuterons tout de suite de la validité de ce construit, en utilisant la définition de « construit » en tant que la meilleure conception d'un trouble, en considérant qu'un trouble est un ensemble de symptômes. La validité de construit se définit donc comme une forme de validité – un processus de mesure de concordance entre deux éléments – d'un

ensemble de symptôme. Nous reformulons d'ailleurs ici la définition de la validité de construit fournie par Cronbach et Meehl en 1955, qui la présentaient comme la validité d'un « réseau » (nomologique)⁶⁰. Elle a été introduite dans un rapport technique de l'*American Psychological Association* en 1954, présidée par Lee Cronbach. L'idée de la validité de construit a été proposée par un des membres de ce groupe de réflexion, Paul Meehl, qui proposait de considérer la validité d'un test comme analogue à la validité d'une théorie scientifique (Cronbach & Meehl, 1955). La validité permettrait de confirmer les prédictions fondées sur la théorie. Afin qu'un construit puisse constituer une catégorie psychiatrique valide, il fallait donc une « théorie », ce qui est en psychopathologie a été traduit de manière pragmatique par l'élaboration d'un ensemble de critères réguliers et formels. Robins et Guze, dès 1970, ont alors proposé que les classifications psychiatriques soient constituées d'un ensemble de critères formels valides (Robins & Guze, 1970). Avec cette proposition, ils offraient aux catégories du DSM une ébauche de validité de construit. Guze avait de plus établi un pont conceptuel entre différentes formes de validité (Guze, 1967, 1975) :

- La validité de construit de la psychologie clinique (Cronbach, 1955) ;
- La validité de structure externe (Hempel, 1959) qui indique que le construit est cohérent avec les autres éléments d'un champ scientifique ;
- La validité de structure interne qui indique que les éléments d'un construit (par exemple, « l'anxiété ») sont *cohérents* entre eux, à l'intérieur du construit.

Le principal construit psychiatrique qui a cherché à répondre à ces trois types de validité correspond à la catégorie retrouvée dans le DSM. Autrement dit, la classification DSM tend à être présentée comme la réunion de la validité de construit, de la validité de structure externe et de la validité de structure interne

⁶⁰ Le terme de « réseau » apparaît exactement 40 fois dans l'article, notamment pour désigner les ensembles de théories et lois qui permettent la définition du construit. Ces lois relient notamment des propriétés ou des quantités observables les unes aux autres. Ainsi, sans discuter de cette notion de loi mais bien en accord avec la définition de Cronbach et Meehl, nous définissons un construit comme un réseau d'éléments (et notamment des symptômes, comme nous le verrons dans la seconde partie de la thèse) en interactions mutuelles.

(Spitzer et al., 1978). Prenons l'exemple du DSM pour mieux comprendre ce qu'il en est des RDoC. Concernant le premier type de validité cité ci-dessus, le terme de « construit » renvoyait selon Cronbach et Meehl à un « réseau » de traits ne pouvant être approchés directement. Les auteurs soulignent eux-mêmes que ces notions proviennent d'une littérature importante qui a précédé leur article, bien que leurs références citées restent relativement pauvres en ce qui concerne le concept de « construit ». Comme le notent Cronbach et Meehl, un construit n'est pas opérationnalisable. En effet, en prenant l'exemple de Hempel (1965) avec les minéraux, il existe une opération possible qui rend valide une définition comme le fait qu'un minéral est plus dur qu'un autre (il s'agit de les frotter l'un à l'autre) (Cronbach & Meehl, 1955). Dans le cadre du DSM, la tentative de validation des construits que sont les catégories est effectuée grâce aux validateurs externes (donc selon une validité de structure externe), qui attestent d'une cohérence par des rapprochements (statistiques) entre les construits et des éléments extérieurs (par exemple, des facteurs génétiques). Brièvement, cette notion de cohérence (mutuelle) signifierait que les relations entre les symptômes d'une catégorie ou les données neuroscientifiques incluses dans la matrice sont le reflet d'une même étiologie sous-jacente (supposée pour les catégories du DSM) ou d'une même fonction cérébrale (située sur une des lignes de la matrice des RDoC). Cette cohérence dépend de la structure interne de ces catégories, mais également de leur contenu, renvoyant à la validité de contenu que nous étudierons dans la seconde partie de la thèse.

Il n'en est pas de même pour les RDoC. Concernant la validité de construit, elle est plus que critiquable : les RDoC s'appuient sur une validité de construits psychologiques déterminés historiquement sur avis d'experts (par exemple, la motivation d'approche, les systèmes cognitifs ou la mémoire de travail). Quand bien même les éléments neurobiologiques intégrés dans la matrice seraient issus de paradigmes expérimentaux solides, les construits (ici, les fonctions) qui les structurent sont potentiellement critiquables. Les RDoC peuvent cependant bien revendiquer une bonne validité de structure interne, car ils intègrent de manière cohérente des niveaux d'analyse issues de leur propre champ disciplinaire, les

neurosciences. Cependant, les RDoC ne peuvent revendiquer posséder une validité de structure externe. En effet, ils n'utilisent volontairement pas de validateurs externes. Les fonctions de la matrice ne renvoient à aucun autre domaine scientifique (autre que neuroscientifique) qui permettrait de les justifier. Phillips écrivait à ce sujet que « la validité des RDoC est vraiment une forme de validité diagnostique forte [telle que décrite par Kendell et Jablensky] comme étant basée sur la définition de caractéristiques fondamentales » (Jablensky et al., 2011). La « validité diagnostique forte » des RDoC s'appuyant sur la « définition de caractéristiques fondamentales » renvoie à la validité de structure interne. Il n'est pas question ici de rechercher une validité externe (rappelons que la validité externe donne sa cohérence à un construit au regard des autres disciplines scientifiques). Cette absence de considération pour la validité de structure externe serait-elle volontaire ? Nous faisons l'hypothèse que la validité externe des RDoC tente d'être justifiée par le fait que les éléments sont empiriquement issus d'études expérimentales. L'espoir des RDoC est que ces éléments puissent être suffisamment cohérents en étant mutuellement intégratifs les uns les autres⁶¹. Autrement dit, nous proposons que le RDoC peut être décrit comme valide en l'absence de toute validité externe explicite⁶² parce qu'il existe une intrication des « lignes » et des « colonnes » de la matrice, qui ne pourraient exister les unes sans les autres. Le projet ne nécessite donc pas de convoquer des validateurs externes. Cependant, il est difficile d'exclure totalement un champ scientifique de son contexte – comme semble le faire les RDoC en ne cherchant pas à satisfaire un certain niveau de validité externe. Prenons par exemple l'exemple de l'influence de la culture sur les construits. Sur le plan culturel, Laurence Kirmayer, qui travaille à l'Université de McGill sur les relations entre santé mentale et culture, propose un cadre de travail dit « écosocial » (Filipe et al., 2020). Celui-ci intègre des validateurs culturels⁶³. De tels validateurs culturels permettent, selon Kirmayer, de renforcer les

⁶¹ Il est important de ne pas confondre les construits qui composent la matrice (vide), que nous discutons ici, des éléments neuroscientifiques dont l'intégration au sein de la matrice dépend des expérimentations neuroscientifiques dont elles sont issues.

⁶² Pour faire un parallèle avec le DSM, notons qu'une catégorie descriptive (par exemple, le trouble obsessionnel compulsif) ne nécessite pas (fort intuitivement) la présence d'une autre catégorie descriptive (par exemple, le trouble anxieux généralisé) pour être validé, ce qui pose d'ailleurs des difficultés en termes de comorbidité.

⁶³ D'ailleurs, dans le chapitre sur les syndromes liés à la culture, coordonné par Kirmayer lui-même, le DSM-5 propose une définition du construit comme « formulation culturelle ».

construits en en expliquant le sens au prisme de l'influence de la culture. Et en allant plus loin encore, la psychiatrie computationnelle pouvant créer des construits intégrant l'aspect culturel (voir par exemple la théorie de la *Thinking Through Other Minds*, amplement discutée dans *Behavioral and Brain Sciences*, 2020 – (Veissière et al., 2020)), les liens d'association entre facteurs physiologiques, sociaux et culturels sont plus clairs. C'est d'ailleurs dans ce sens que les construits peuvent être définis comme les résultats de « réseaux » (nomologiques).

Pour résumer, le DSM et les RDoC s'appuient tous deux sur une validité de construit. En plus de celle-ci, le DSM s'appuie majoritairement sur une validité de structure externe via des validateurs externes, qui relie ses catégories à des éléments extérieurs à la classification. Les RDoC reposent principalement sur une validité de structure interne, qui renvoie à la cohérence des domaines de fonctionnement entre eux. Cette validité de structure interne semble alors s'appuyer sur l'intégration mutuelle des différentes entités des RDoC (les domaines en ligne et les unités d'analyse en colonne). En cela, les RDoC ont cherché à éviter de justifier une quelconque validité externe. Afin de mieux comprendre comment les différents éléments s'intègrent entre et avec les domaines de recherche scientifique qui entourent la matrice, la section suivante portera sur la notion d'intégration.

IV. Les RDoC et les trois niveaux d'intégration

L'enjeu de l'intégration entre les différents éléments des RDoC (que nous allons définir) est majeur pour comprendre notre problématique, à savoir de comprendre si les RDoC soulèvent des difficultés similaires au DSM du fait de leur structuration sur une approche par la dimension / catégorie. Ici, nous faisons l'hypothèse que c'est en partie au travers de cet espoir d'intégration entre éléments de la matrice que les RDoC justifient l'existence des fonctions (ou construits, ou dimensions). Le projet RDoC propose des « variables *indépendantes* étudiées dans différents domaines expérimentaux » (nous soulignons)

(Cuthbert, 2019), organisées autour de fonctions (par exemple, la peur, la mémoire de travail, le retrait social). Cependant, cette « indépendance » des variables s'applique aux unités d'analyse en tant qu'elles s'inscrivent, chacune, dans « une discipline scientifique » à part entière. En effet, l'unité d'analyse des « gènes » (qui a été supprimée de la matrice secondairement) constitue une discipline scientifique qui n'utilise ni les concepts, ni la sémantique, ni les modèles, ni la théorie de l'unité d'analyse adjacente, « molécules ». Alors comment la matrice des RDoC peut-elle faciliter le développement « d'explications psychobiologiques intégratives » comme cela a été soutenu par exemple par Cuthbert et Kozak en 2013 (Cuthbert & Kozak, 2013) et à de multiples reprises par la suite ? Cette question est d'autant plus importante que nous faisons l'hypothèse qu'elle conditionne la validité de structure interne du projet (qui serait fondée sur l'intégrativité des éléments de la matrice entre eux). La question de l'intégration a soulevé de multiples débats en philosophie des sciences. Par exemple, des auteurs comme Brigandt (Brigandt, 2010, 2010, 2013b) ont soutenu l'existence de différents types d'intégration. Nous choisissons de définir l'intégration de manière très large comme un processus qui « combine une variété d'idées différentes [issues] de différents domaines scientifiques » et plus particulièrement la possibilité que « des connaissances appartiennent à différents niveaux d'organisation » (Brigandt, 2010).

Afin de mieux comprendre comment cet enjeu d'intégration cherche à justifier l'existence des fonctions, reliant le projet à une approche par la dimension, nous proposons de décrire trois types d'intégration en neurosciences et pour la psychiatrie. Au sein du cas concret que représente le projet RDoC, nous décrivons un type d'intégration qui concerne ses dimensions (ou fonctions), un type d'intégration qui concerne la tentative d'application des RDoC à la clinique psychiatrique et un type d'intégration « interdisciplinaire ». Les deux derniers type sont essentiels car ils permettent de comprendre comment les RDoC s'insèrent dans les communautés cliniques et scientifiques, au-delà de la simple compréhension de la structure de la matrice.

1. Intégrations mutuelles entre les principales fonctions cérébrales (entre les lignes de la matrice)

Détailler le potentiel d'intégrativité mutuelle des fonctions (ou dimensions, ou construits) de la matrice des RDoC permet de mieux comprendre en quoi le projet constitue fondamentalement une approche par la dimension (qui, nous l'avons dit en introduction, comporte des enjeux superposables à l'approche par la catégorie). Nous avons vu précédemment que les fonctions étaient incluses dans la matrice (par les groupes d'experts) si elles avaient été validées par des études préalables, c'est-à-dire s'il existait des données probantes qui confirmaient leur capacité d'expliquer un phénomène comportemental. Ces dimensions sont donc réunies au sein de la matrice des RDoC car elles comportent toutes une certaine « validité ». Ce constat pose donc la question de la *validité des résultats des études expérimentales*, que nous allons discuter. Les études scientifiques qui utilisent le cadre RDoC peuvent s'appuyer sur *plusieurs* de ces fonctions. Par exemple, elles peuvent utiliser la fonction de « frustration non récompensée », appartenant au système de valence négative, en même temps que la « motivation à l'approche », appartenant au système de valence positive. De nombreux articles utilisant le cadre RDoC intègrent ainsi plusieurs fonctions (Kozak & Cuthbert, 2016), même si d'autres sont moins intégratifs (Insel et al., 2010). La **Figure 4**, extraite d'un article portant sur l'intégration des troubles de personnalité au sein de plusieurs dimensions (mais pas au sein des unités d'analyse) de la matrice RDoC, en est un exemple (Koudys et al., 2019). Dans ce tableau, un trouble de personnalité paranoïde (deuxième colonne) selon le DSM peut être décliné en plusieurs dysfonctions selon le cadre RDoC (sept fonctions : menace potentielle et soutenue, attention, affiliation et attachement, communication sociale, perception et compréhension d'autrui et éveil). S'il n'existe pas d'intégration entre fonctions au sein de la matrice elle-même, puisque chaque donnée (par exemple, un mécanisme ou une molécule) neuroscientifique n'appartient qu'à une « case » de la matrice, cette intégration peut se faire par l'intermédiaire d'un regroupement de fonctions pour un phénotype clinique particulier, comme un trouble de la personnalité.

Notons cependant que ces phénotypes particuliers sont majoritairement sous-tendus par les catégories du DSM (comme cela est le cas dans la **Figure 4** au sujet des troubles de personnalité).

Table 1 Preliminary conceptual mapping of NIMH research domain criteria constructs onto DSM-5 personality disorder criteria

Construct	PAR	SCZ	SZT	ASPD	BPD	HIS	NPD	APD	DEP	OCP
Negative Valence Systems										
Acute Threat ("Fear")		6		4, 5	1, 8, 9			4	6, 7	
Potential Threat ("Anxiety")	1, 3, 4, 7		5, 9	6	5			1, 2, 3, 4, 5, 7	3, 8	
Sustained Threat	1, 2, 5, 7		9					4		7
Loss					7			4	7	
Frustrative Nonreward					8	1	2, 4, 5, 8			
Positive Valence Systems										
Reward Responsiveness		1, 4, 6			2, 7					
Reward Learning			2		2	2	1, 5	2		2, 3, 4, 8
Reward Valuation		1, 2, 3	8	2, 3, 5	2, 4	1, 4, 8	1, 2, 3, 4, 6, 8	2	5, 7	3, 5, 7
Cognitive Systems										
Attention	4, 7		1							1
Perception			3		9					
Declarative Memory										
Language			4							
Cognitive Control				3, 6	3, 4, 8	6			1, 2, 4	1, 2, 4, 6, 8
Working Memory										
Social Processes										
Affiliation and Attachment	3, 5, 6	1, 2, 3, 5, 7	8	1, 2, 4, 5, 7	1, 2, 8	1, 2, 4, 7, 8	4, 6, 7, 9	1, 2, 3, 4, 5	2, 3, 5, 7, 8	3, 6, 8
Social Communication	4	6, 7	4, 6, 7	7	1	2, 3, 4, 5, 6		4		
Perception and Understanding of Self			1, 2, 3		3, 8, 9		1, 2, 3, 4, 5, 8	5, 6, 7	1, 2, 4, 6, 7, 8	4
Perception and Understanding of Others	1, 2, 3, 4, 6, 7	6	5, 9	7	1, 2, 8, 9	7, 8	1, 3, 4, 5, 6, 7, 8	1, 2, 3, 4, 7	3	1, 4
Arousal and Regulatory Systems										
Arousal	6	4, 6, 7	9	4, 7	1, 5, 6, 7, 8, 9	1	7	4, 5	6	5, 6
Circadian Rhythms										
Sleep-Wakefulness										

Each number references the diagnostic criterion for the respective DSM-5 personality disorder

PAR paranoid, SCZ schizoid, SZT schizotypal, ASPD antisocial, BPD borderline, HIS histrionic, NPD narcissistic, APD avoidant, DEP dependent, OCP obsessive-compulsive

Figure 4. Les construits diagnostics du DSM-5 (première ligne), correspondant aux catégories DSM des troubles de personnalité, sont mis en relation avec les principales fonctions cérébrales implémentées dans les RDoC. Les numéros renvoient au critère du trouble de personnalité correspondant. Issu de (Koudys et al., 2019) (site du NIMH).

Pour résumer, plutôt que de tenter de traiter un unique trouble psychiatre selon une catégorie traditionnelle du DSM, un large éventail de ces troubles peut être traité par un unique protocole de recherche, qui s'appuie sur plusieurs dimensions (ou fonctions) de la matrice RDoC. Le fait de traiter dans un même protocole de recherche plusieurs diagnostics traditionnels issus du DSM peut être qualifié d'*approche transdiagnostique* (Krueger & Eaton, 2015; Mansell, 2019; McGorry & Nelson, 2019; Monestès & Baeyens, 2016). En cela, les RDoC cherchent à s'inscrire dans une approche

translationnelle. Dans ce cas, la volonté de « translation⁶⁴ » est réalisée entre la recherche (un ensemble de fonctions cérébrales, par exemple les sept fonctions citées précédemment) et la clinique (un comportement phénotypique, par exemple le trouble de personnalité paranoïde). La science fondamentale (celle des fonctions cérébrales) est en quelque sorte « traduite » en comportements cliniquement observables – ou autrement dit, les troubles sont pensés en termes de mécanismes cérébraux dysfonctionnels.

De fait, en ce qui concerne le premier type d'intégration entre fonctions de la matrice, nous affirmons que l'intégration mutuelle est partiellement réalisée dans le système RDoC. L'intégration de plusieurs fonctions conjointes, dans une unique étude, peut théoriquement être réalisée. Cependant, les RDoC ne permettent pas qu'une fonction soit intégrée à une autre : chaque fonction est considérée comme une « boîte » isolée d'une autre. En cela, le projet se rapproche fortement d'une approche par la catégorie⁶⁵.

2. Intégrations mutuelles entre unités d'analyse (entre les colonnes de la matrice)

Peut-on espérer qu'une intégration entre unités d'analyse permette de mieux justifier l'existence des fonctions, soutenues par une telle approche par la dimension ? Le projet RDoC est théoriquement constitué d'une matrice vide que les études expérimentales doivent « remplir ». Autrement dit, les RDoC

⁶⁴ Ou « traduction », mais ce terme n'est volontairement pas utilisé (en anglais comme en français) pour éviter d'avoir à justifier la transformation qu'un construit ou un résultat subirait par suite de son transfert d'une matrice disciplinaire (orientations, intérêts, perspectives et méthodes d'un domaine scientifique) à une autre. La traduction pourrait en effet être sémantique, conceptuelle, sociologique, etc. Ici, le terme de 'translation' renvoie plutôt à l'idée du « glissement » du construit, identifié en recherche, vers la clinique.

⁶⁵ De plus, le projet peine à être pluraliste. En effet, ce n'est pas parce que plusieurs fonctions sont utilisées « en même temps » ou « chez le même individu » que nous pouvons parler de pluralisme explicatif. En même temps, cette même constatation peut être tempérée en soutenant que les RDoC sont pluralistes *si l'on se place du point de vue uniquement neuroscientifique*. Les neurosciences étant constituée d'un ensemble de disciplines, de méthodes, de chercheurs et de laboratoires différents, le projet RDoC leur offre un cadre de travail commun. Inversement, si l'on considère le projet RDoC d'un point de vue multidisciplinaire, il apparaît bel et bien réductionniste. Il est alors réductionniste dans la mesure où il n'est pas capable d'inclure directement des paramètres qui ne sont pas neuroscientifiques, par exemple des paramètres sociologiques ou culturels.

ne proposent qu'un ensemble de critères (les *research domain criteria*) qui doivent être complétés par des données neuroscientifiques (autrement dit, il faut « remplir les cases »). Ce point de départ théorique est différent de celui des critères de Feighner, des RDC ou du DSM. En effet, ceux-ci incluaient à la fois des critères (par exemple, critère A., B., C., etc.), qui répondaient à l'existence de données cliniques (par exemple, tristesse de l'humeur, hallucinations, etc.), issues certes de l'histoire de la psychiatrie, mais bien « déjà-là », sous la main des cliniciens et concepteurs du DSM. Dans les RDoC, les données neuroscientifiques devraient être incluses dynamiquement, sans cesse renouvelées par la recherche en neurosciences (à la différence des critères de Feighner traditionnels qui relevaient de l'inscription dans la tradition clinique psychiatrique descriptive). Ainsi, nous avançons que les RDoC constituent un « réceptacle à données » voué à être sans cesse complété. Le projet s'inscrit donc dans le temps, du fait même de sa nécessité d'être complété par itérations successives⁶⁶. Ces données sont recueillies à différentes échelles du vivant et implémentées dans la matrice dans chaque unité d'analyse. Mais comment peuvent-elles être intégrées les unes aux autres ?

La matrice des RDoC recense des variables considérées comme des « mécanismes neuroscientifiques » qui pourraient être des « cibles pour améliorer les diagnostics ou les interventions thérapeutiques ». Pour comprendre le sens de la clinique et pragmatique de chaque fonction, ces mécanismes (à l'intersection des lignes et des colonnes) devraient pouvoir être intégrés les uns aux autres, c'est-à-dire reliés par un processus général ou régional, non seulement en ce qui concerne les niveaux d'organisation (les colonnes), mais également en ce qui concerne les construits (les lignes). Autrement dit, la donnée neuroscientifique située dans « une case » de la matrice des RDoC (à l'intersection d'une ligne et d'une colonne) devrait pouvoir être mise en relation avec la donnée d'une autre « case » de la même ligne (ou même fonction). Par exemple, au sein du système de valence négative (en ligne), on retrouve la

⁶⁶ Tout en notant que ce « réceptacle » est lui-même régulièrement mis à jour, par exemple avec l'ajout d'un sixième domaine sensorimoteur.

dimension « menace potentielle ('anxiété') »⁶⁷. Dans la colonne « molécules » correspondante, au sein de la « case », on retrouve un élément issu d'une donnée neuroscientifique, par exemple le cortisol. Si l'on prend la colonne suivante, « cellules » pour la même fonction « menace potentielle ('anxiété') », on retrouve dans la « case » les cellules pituitaires. Or, quel rapport peut-il y avoir entre le cortisol et les cellules pituitaires, qui pourtant renseignent toutes deux sur la fonction de menace potentielle ? Sur le plan purement biologique, il n'y a pas de lien entre le cortisol et la glande pituitaire. Mais surtout, sur le plan épistémologique, rien ne permet de rendre compte d'une quelconque relation entre ces « cases ». À l'instar de l'intégration entre fonctions, nous avançons que l'intégration des unités d'analyse est manquée par les RDoC. Comme dans le DSM, aucun élément structurel du projet RDoC ne permet d'établir une relation entre ses unités d'analyse.

Pourtant, ses promoteurs vantent les mérites d'intégration des RDoC. Cuthbert écrivait dès 2014 : « Un important projet de recherche des RDoC sera de se concentrer sur un problème clinique spécifique, qui pourrait être mieux expliqué grâce à un protocole de recherche *combinant* différentes unités d'analyse, dans le but d'éclairer les mécanismes de dysfonctionnement mis en jeu » (nous soulignons) (Cuthbert, 2014). Afin de bien comprendre comment ces différentes unités d'analyse sont « combinées », nous pourrions décrire les RDoC comme relevant des deux premiers niveaux de complexité décrits par Bechtel et Richardson : ils possèdent une complexité de composants et une complexité agrégative (Bechtel & Richardson, 2010). À la manière des pièces d'une horloge, chaque élément est indépendant (complexité de composant), bien que chacun soit nécessaire pour la fonctionnalité de l'horloge (complexité agrégative). Cependant, les RDoC ne relèveraient pourtant pas le défi d'un troisième niveau de complexité, qualifié d'intégré, dans lequel les pièces de l'horloge perdent leur indépendance et ne peuvent être étudiées qu'en tant que composants du système⁶⁸. Les RDoC ne permettraient pas

⁶⁷ Concernant cette dimension, on note d'ailleurs qu'elle renvoie implicitement à la catégorie classique « d'anxiété » sans pour autant chercher à s'y conformer absolument (et en la laissant d'ailleurs entre parenthèses puisqu'il s'agit uniquement d'un « label » qui n'est qu'un construit sans correspondance clinique)

⁶⁸ Ce qui correspondant au fait que « le tout est plus que l'ensemble des parties », comme décrit en systémique ou dans les thèses holistes.

d'expliquer la complexité intégrée, dans laquelle chaque mécanisme ne pourrait être étudié qu'en tant que composant d'un système général (Bechtel & Richardson, 2010). Autrement dit, les fonctions des RDoC non intégrées les unes aux autres, décrites par des mécanismes non intégrés les uns aux autres et à des échelles du vivant non intégrées les unes aux autres ne permettraient pas de comprendre les symptômes. Ils permettraient seulement d'en décrire les caractéristiques (Bechtel, 2008). L'échec de l'intégration entre fonctions et unités d'analyse au sein de la matrice des RDoC réduit le projet à une approche pas la catégorie tout aussi classique et traditionnelle que le DSM. Par extrapolation, nous pourrions avancer que tout projet neuroscientifique construit sur une approche par la catégorie ou la dimension peinera à s'extraire d'une vision traditionnelle de la psychiatrie telle que la proposait le DSM. Nous décrirons dans la section suivante que la promesse d'intégrativité pourrait cependant être pensée grâce à une intégration de la matrice « avec » et « dans » la coordination entre différentes disciplines (neuro)scientifiques. Nous verrons dans la seconde partie de cette thèse comment s'extraire de cette approche par la catégorie et dimension.

3. Intégration sur le plan interdisciplinaire

Après avoir cherché à comprendre comment pouvait se réaliser l'intégration mutuelle entre fonctions (les lignes) ou entre unités d'analyse (les colonnes), nous considérons désormais qu'il ne peut y avoir d'intégration au sein de la matrice des RDoC. La matrice ne constitue pas en elle-même une approche intégrative. Nous détaillerons cependant comment cette intégration se réalise *au-delà* de la matrice elle-même, grâce à son application pratique en recherche et à un niveau social.

Les RDoC ont été présentés comme une feuille de route pour développer la recherche. En effet, la matrice des RDoC sert de base de travail pour synthétiser un ensemble d'éléments biologiques et permet une meilleure collaboration entre des chercheurs appartenant à différents domaines neuroscientifiques. Nous

avançons que c'est grâce à une telle collaboration que l'intégration des différentes dimensions peut être réalisée. En affirmant cela, nous adoptons donc la conception de l'intégration de Brigandt (2010) considérant l'intégration comme le processus de combinaison de différentes connaissances appartenant à différents domaines scientifiques. L'intégration ne serait donc pas dépendante des lignes ou des colonnes de la matrice, mais proviendrait de leur utilisation au sein de différentes disciplines et laboratoires de recherche. Par exemple, un laboratoire spécialisé dans la manipulation des cellules (niveau cellulaire) travaillerait avec un laboratoire expert en imagerie fonctionnelle (niveau de la connectivité) autour d'une même fonction (et souvent, comme nous l'avons vu, de plusieurs fonctions) appartenant au circuit de la récompense. Dans notre hypothèse, nous revenons d'une certaine manière à l'idée de (Grüne-Yanoff, 2016) d'une « interdisciplinarité sans intégration ». Celui-ci soutient que la notion d'intégration n'a pas de valeur épistémique en soi, mais que c'est à travers l'interdisciplinarité que différents objets scientifiques (ici, les fonctions et les unités d'analyse de la matrice) peuvent être travaillés ensemble. Nous soutenons plutôt ici que la notion d'intégration peut être pertinente en admettant la nécessité d'une interdisciplinarité. Autrement dit, c'est au travers du travail réciproque et mutuel entre différentes communautés cliniques et de recherche, par exemple au sein d'un protocole d'étude commun, qu'une intégration entre fonctions peut être réalisée. Une telle intégration nécessite de considérer à la fois des niveaux sociologique (acteurs, communautés, controverses, points de passage obligés, objets-frontières, etc.) et économiques ou politiques (financements inter-laboratoires, propositions de projets académiques transversaux, dynamiques des processus technoscientifiques, etc.).

4. Conclusion sur l'intégration

Pour conclure l'ensemble de cette section portant sur l'intégration, nous soutenons que la matrice des RDoC, à elle seule, ne peut permettre ni l'intégration mutuelle de ses fonctions, ni celle de ses unités

d'analyse. Cette incapacité renforce l'idée que les différents éléments des RDoC ne peuvent être considérés indépendamment les uns des autres. De fait, nous pourrions dire que le projet lui-même reste une « coquille vide neuroscientifique », constitué d'un ensemble de « boîtes fonctionnelles », surtout capables d'accueillir des tests expérimentaux. Du fait de cette structuration, le projet s'apparente bien à une approche par la dimension / catégorie qui véhicule les mêmes limites que celle que nous avons critiquée pour le DSM. En considérant l'utilisation de la matrice de manière interdisciplinaire, il semble cependant y avoir une forme d'intégration qui est rendue possible par le travail conjoint entre individus de communautés et d'institutions de recherche différentes. Au lieu de proposer une interprétation universelle de ce qu'est l'intégration, dans le cas des RDoC, nous proposons que l'intégration soit d'abord amenée sur un plan sociologique par l'association des sciences du cerveau avec les autres champs disciplinaires. Autrement dit, nous soutenons que la meilleure intégration permise par le projet RDoC se situe justement au travers de sa capacité à faire interagir différentes communautés scientifiques, amenant discussions et controverses (Star & Griesemer, 1989). Cependant, l'application des résultats d'une discipline neuroscientifique à la pratique clinique relève d'un travail épistémologique plus profond qui ne semble pas pouvoir être relevé par les RDoC, comme nous allons le détailler.

V. RDoC, intégration du neurodéveloppement et application clinique

Dans une perspective d'applicabilité à la clinique, le projet a donc fait le pari d'une validité de ses fonctions, en promettant une intégration entre elles. Autrement dit, l'initiative de validité et la volonté d'intégrativité auraient pour objectif une meilleure applicabilité en pratique. Comment les RDoC se positionnent-ils réellement par rapport à la pratique clinique ? Pour répondre à cette interrogation, nous questionnerons la manière avec laquelle les RDoC parviennent à mobiliser la part contextuelle et

environnementale qui constitue intrinsèquement les troubles psychiatriques. Dans cette dernière section, nous présenterons la manière avec laquelle le projet RDoC modélise les comportements et, à travers la thématique du neurodéveloppement, comment y est conçu l'environnement. Nous montrerons ainsi que les RDoC possèdent une valence clinique, même si l'alliance entre RDoC et clinique pose autant d'interrogations qu'elle ne parvient à en résoudre.

1. RDoC et comportements

Les unités d'analyse pourraient être potentiellement reliées aux comportements. Du fait de la complexité des troubles mentaux, l'histoire de la psychiatrie de ce dernier siècle a montré qu'il n'était pas possible d'affirmer que, sur un simple test d'imagerie, un tel patient serait souffrant, par exemple, de schizophrénie (Kendler, 2016). Le contexte, l'environnement ou le ressenti expérientiel du sujet sont quelques-uns des facteurs nécessaires pour comprendre le comportement, ou phénotype comportemental. Pour répondre à cette difficulté, les RDoC font l'hypothèse que les mesures cérébrales seraient le reflet des comportements. En d'autres termes, la matrice des RDoC permettrait de traduire en « langage cérébral » les facteurs environnementaux, sociaux et d'exposition. Elle cherche à transcrire les événements extérieurs en perturbations cérébrales en s'appuyant sur différentes (dys)fonctions du cerveau. C'est pour cette raison que le projet RDoC a cherché à distinguer des unités d'analyses (en colonnes), représentant le cerveau à différentes échelles, de fonctions (en lignes), représentant les comportements. Nous pourrions décrire l'étude des comportements menés par les RDoC en décrivant brièvement un de ses projets, le projet européen PRISM (pour *Psychiatric Ratings Using Intermediate Stratified Markers*). Celui-ci cherche à explorer les aspects du comportement social et de la cognition sociale (Cuthbert, 2019) en s'appuyant sur le cadre RDoC et plus particulièrement sur la dimension des processus sociaux. Le projet PRISM est un programme de recherche financé par la *European Union's*

Innovative Medicines Initiative (IMI) (Shah et al., 2020). Son principe fondamental est d'identifier des marqueurs sociaux et de les relier aux déterminants biologiques. Il s'agit donc de produire des études qui cherchent à relier des mesures biologiques avec des mesures comportementales. Ces études mettent en relation des marqueurs *cérébraux* impliqués dans des comportements *sociaux*. Par exemple, dans le cadre de la schizophrénie, une des possibilités offertes par les RDoC est de décomposer les comportements « extérieurs » de la personne en différents niveaux d'analyse : troubles du langage, hallucinations visuelles, troubles du fonctionnement cognitif, difficulté d'organisation ou de motivation. Pour chaque groupe de patients, cette décomposition des comportements est « codée » dans un modèle génétique (par exemple, le syndrome de Williams-Beuren est dû à une microdélétion chromosomique située dans la région 11q.23 d'un des deux chromosomes 7 impliquant la dimension du processus social de la matrice RDoC), une modification moléculaire et cellulaires (par exemple, l'acétylcholine est une molécule pouvait être liée au comportement social de stress impliquant le système de valence négative de la matrice RDoC), une activité des circuits de la connectivité cérébrale (par exemple, les états de dissociation ou d'hypervigilance peuvent impliquer le réseau du mode par défaut, impliquant le construit de l'attention dans le domaine du système cognitif de la matrice), etc. Tous ces niveaux correspondent aux unités d'analyse de la matrice RDoC.

Ainsi, les unités d'analyse sont mises en relation avec les comportements. Les RDoC proposent de traduire les comportements en fonctions cérébrales. À l'inverse, des perturbations dans le système de reconnaissance des émotions (RDoC) se « visualisent » cliniquement à travers des difficultés à répondre à une situation sociale complexe (comportement). Cette traduction des événements extérieurs en éléments biologiques est à la base de la légitimité du projet RDoC. Cette notion de légitimité, touchant à la traduction entre neurosciences et comportements, soulève des questions fondamentales auxquelles les RDoC ne peuvent pas systématiquement répondre. Par exemple, peut-on imaginer des indicateurs environnementaux qui ne pourraient pas être « codés » dans un langage cérébral ? De plus, les altérations des fonctions (par exemple cognitives, comme la mémoire de travail, la planification ou la

reconnaissance des émotions) sont présentes à la fois chez des patients atteints de schizophrénie, de troubles dépressifs, de troubles bipolaires ou dans la population générale : la traduction des éléments extérieurs en éléments biologiques concerne-t-elle tout individu, qu'il soit ou non considéré comme porteur d'un trouble psychiatrique ? Les RDoC seraient-ils condamnés à mettre de côté la distinction entre normal et pathologique au prix d'une traduction comportementale ? Nous allons voir que ces apparentes limites pourrait être nuancées.

2. RDoC et neurodéveloppement

Le cerveau n'est pas un organe statique. Sa dynamique est largement influencée par le développement ainsi que par l'expérience. Selon cette perspective, le cerveau « mûrit », avec des périodes critiques, en fonction des interactions avec les événements extérieurs. Une lecture superficielle de la matrice pourrait laisser penser que les dynamiques de développement ne sont pas concevables au sein des RDoC. En effet, la dynamique évolutive des fonctions cérébrales, nécessaires à la compréhension du neurodéveloppement – dynamique par essence – ne semble pas incluse au sein de la matrice. Nous soutenons cependant dans cette section que les RDoC se positionnent comme un cadre de travail permettant de structurer cette dynamique (Casey et al., 2014; Eijndhoven et al., 2021; Mittal & Wakschlag, 2017).

Neuroscientifiques, généticiens, épidémiologistes et psychiatres cliniciens sont constamment confrontés à des perspectives développementales. Par exemple, l'épidémiologie et la génétique de la dépression ne sont pas les mêmes dans l'enfance, l'adolescence ou la vie adulte (Hauser et al., 2018). La mise en évidence de telles différences soulève la question de savoir, par exemple, si la communication sociale présente chez un patient porteur d'un trouble du spectre de l'autisme a la même signification clinique à différents âges. De même, l'irritabilité ne peut pas être étudiée de la même manière chez un jeune enfant

que chez un sujet adulte autonome – et ce qui est nommé « irritabilité » n’a pas la même valeur diagnostique et thérapeutique en fonction des contextes cliniques. L’approche des RDoC propose justement qu’une de ses dimensions (par exemple le système de valence négative) puisse s’adapter à toutes les périodes de la vie (Casey et al., 2014). Par exemple, l’élément de menace (ligne) au niveau physiologique (colonne) du système de valence négative (fonction, ou dimension), à savoir la dysrégulation de l’axe hypothalamo-hypophysaire, aurait une portée explicative identique pour un comportement du petit enfant que chez l’adolescent, mais pas chez l’adulte. Autrement dit, les dimensions des RDoC intègrent la composante temporelle et évolutive du patient.

Expliquons comment les RDoC permettent de comprendre le neurodéveloppement. Les principaux troubles neurodéveloppementaux décrits au sein du DSM-5 comprennent l’autisme, le trouble déficit de l’attention avec ou sans hyperactivité, les troubles des apprentissages, les troubles de communication (pragmatique) et la déficience intellectuelle. Cependant, une importante littérature soutient que les autres troubles, comme la bipolarité ou la schizophrénie, pourraient avoir également une forte composante neurodéveloppementale (Owen et al., 2011; Riglin et al., 2019). L’idée de concevoir les troubles psychiatriques selon une perspective développementale est loin d’être nouvelle (Harris, 2014). Cependant, l’arrivée de nouvelles techniques génétiques, comme les *Genome Wide Association Studies* (Power et al., 2017) ou méthodologiques, comme les modèles computationnels, permet de repousser les limites de la compréhension des troubles psychiatriques. Par exemple, ces techniques permettent de rassembler des phénotypes différents (par exemple, par des études génétiques), de prédire des trajectoires d’évolution des troubles (par exemple, par l’intelligence artificielle) ou de stratifier en sous-groupes (par exemple, par des analyses de profils immunologiques ou moléculaires). Ces progrès montrent l’importance de considérer la dynamique des troubles, comme cela est concrétisé à travers différents modèles (staging, modèle 3P pour « prédisposition, précipitation, perpétuation », systèmes dynamiques, etc.). Ces techniques et modélisations tendent à s’aligner sur les nécessités de la clinique. Par exemple, un service de soin psychiatrique qui cherchera à recueillir des informations sur de potentielles anomalies

généétiques sous-jacentes à un trouble devra interroger la famille du patient afin de créer un arbre généalogique et devra s'attarder sur les anomalies précoces pouvant faire évoquer une quelconque anomalie associée soutenant l'hypothèse d'une mutation génétique (par exemple, un retard dans les acquisitions psychomotrices ou des troubles neurologiques associés, à l'image de ce qui est fait dans les consultations de neurologie). De même, un service de thérapies intégratives qui chercherait à trouver la meilleure approche psychothérapeutique pour un patient donné devra s'enquérir de facteurs neurodéveloppementaux propres au patient, par exemple, en analysant des schémas de réponse comportementale appris dans la jeune enfance pour développer une thérapie des schémas. Cette importance donnée au neurodéveloppement est concrétisée par les RDoC par la mise en place d'études longitudinales, qui analysent les dynamiques de développement au fil du temps, par exemple, de la grossesse à la puberté. Par exemple, des études prospectives étudiant les sujets à haut-risque (comme les enfants de parents atteints d'une maladie mentale sévère ou l'étude des familles porteuses d'un facteur de risque prédisposant à une maladie) sont intégrés à des bases de données longitudinales évoluant sur des longues périodes de temps (Duffy, 2010). Bien que cela ne soit pas évident à la simple lecture de la matrice, en soutenant de tels projets longitudinaux, les RDoC ancrent la psychiatrie dans une perspective dynamique et neurodéveloppementale. Ce constat montre que le projet peut constituer un soutien aux études sur les phénotypes dans leur composante évolutive⁶⁹. Plus précisément, comme nous allons le détailler ci-dessous, les RDoC cherchent à mêler la notion de neurodéveloppement avec celle d'environnement. Le projet soutient en effet l'analyse des déterminants biologiques du neurodéveloppement, eux-mêmes conséquences de perturbations environnementales. Nous allons détailler ces relations.

⁶⁹ D'ailleurs, certains de ses défenseurs diront que cette idée que les RDoC ne peuvent intégrer de valence dynamique vient d'une contrainte pratique de la matrice : il est simplement difficile de représenter l'espace à quatre dimensions dans une publication scientifique (les trois premières « dimensions » concernent les fonctions cérébrales, les unités d'analyse et les données neuroscientifiques incluses dans chaque case de la matrice), ce qui justifierait le fait que la matrice RDoC ne présente que des lignes et des colonnes. Nous retranscrivons ici l'idée de Sarah Morris, une panéliste de l'intervention de présentation du projet RDoC, intervention consultable sur le site du NIMH, 2015.

3. RDoC et clinique

Jaspers, en 1963, soutenait que les soins cliniques devraient répondre à la demande telle qu'elle est exprimée par le patient (Jaspers, 1963). En effet, ce sont généralement les plaintes liées aux symptômes qui amènent à ce que des soins soient indiqués. Autrement dit, c'est la composante *préjudiciable* décrite au travers des symptômes qui amène à prodiguer des soins. Cependant, du fait de leur prise en compte exclusive des *dysfonctions*, les RDoC s'écartent de cette attention portée à la sémiologie. De ce fait, de nombreux auteurs ont questionné l'utilité clinique des RDoC, comme Maj en 2014 :

« L'écart est bien marqué entre les dimensions proposées par les RDoC et les signes et symptômes qui, comme le reconnaît Cuthbert, constituent "les phénomènes cliniques réels qui amènent les patients à la clinique". Si le problème des catégories du DSM pouvait être qu'elles sont trop éloignées des neurosciences, le problème avec au moins certains des construit des RDoC pourrait être qu'ils sont quelque peu éloignés des phénomènes cliniques » (Maj, 2014)⁷⁰.

Mais pour discuter de l'intérêt clinique des RDoC, nous allons rappeler les apports d'une discipline adjacente à la sémiologie psychiatrique : la neuropsychologie. Partant du pari d'intégrer la neuropsychologie à la pratique clinique psychiatrique sont nés des courants institutionnels (par exemple, l'Association Francophone de Remédiation Cognitive), des pratiques de soins (par exemple, les bilans de cognition sociale), des conceptions et des modèles de la maladie (par exemple, le trouble de l'agentivité dans la schizophrénie). Ce pari intégratif a également permis l'ouverture des services de psychiatrie clinique à de nouvelles professions (les neuropsychologues), conduisant ainsi à de nouvelles thérapies (par exemple, la remédiation cognitive). La neuropsychologie a permis de comprendre

⁷⁰ On pourrait également penser au titre de l'article de Parnas dans le numéro de 2014 de *World Psychiatry* dédié aux RDoC, *The RDoC program: psychiatry without psyche?* (Parnas, 2014).

différemment les symptômes des patients atteints d'un trouble psychiatrique. En effet, cette discipline a mis en avant le fait que les patients porteurs d'un trouble psychiatrique présentaient des difficultés dans la cognition sociale (reconnaissance des émotions, théorie de l'esprit, style attributionnel et perception et connaissance sociales) ou dans la neurocognition (fonctions exécutives, mémoire ou attention). Ces constats neuropsychologiques sont directement applicables à un patient, dans le cadre temporel et spatial d'une consultation tout en ne nécessitant que peu d'équipements ou de technologies.

Cette intégration « réussie » de la neuropsychologie à la clinique psychiatrique pourrait laisser penser que l'approche neurobiologique des RDoC pourrait connaître le même destin (McCoy et al., 2020). Les RDoC auraient été pressentis comme formant un tremplin pour guider la psychiatrie vers une approche plus neuroscientifique (Insel et al., 2010). Il est vrai que les multiples niveaux d'explication fournis par les RDoC contribuent à la « compréhension aidée par l'explication » – pour reprendre les termes de Kendler et Campbell, selon lesquels les niveaux d'explication peuvent profiter aux thérapeutes dans l'éducation thérapeutique des patients (évitant l'explication naïve du déséquilibre chimique) ou dans la compréhension de l'expression des symptômes (Kendler & Campbell, 2009). De même, la connaissance des transformations biologiques du cerveau ne serait pas un obstacle à la pratique clinique. Et grâce à ses multiples niveaux d'explication, le projet RDoC n'empêcherait pas spécialement le développement d'une clinique empathique et tournée vers une meilleure compréhension des symptômes exprimés par les patients. Il renforcerait la valeur clinique en associant à la sémiologie des connaissances neuroscientifiques, dimensionnelles et, d'une certaine manière, agnostiques (c'est-à-dire se dégageant de tout diagnostic traditionnel) (Bolton, 2016). Par exemple, des psychothérapies basées sur les dimensions des RDoC et adaptées aux particularités de chaque patient, contribuent aux avancées thérapeutiques fondées sur ce projet. Ainsi, un programme psychothérapeutique comme « Engage », s'appuyant justement sur les RDoC (Alexopoulos & Arean, 2014), implique d'abord l'identification des fondements neurobiologiques des troubles, ensuite la mise en place d'interventions structurées basées

sur ces fondements et enfin une évaluation de l'efficacité du traitement en relation avec les modifications observées dans ces fondements

Le projet RDoC pourrait donc enrichir plutôt qu'appauvrir la psychiatrie clinique. Cependant, en pratique, un tel apport ne pourrait être valable que si la redistribution des intérêts et des moyens vers cette conception de la psychiatrie n'était pas réalisée au profit d'autres conceptions. Or, les budgets de fonctionnement (par exemple de la division des neurosciences et des sciences fondamentales du comportement du NIMH, la *Division of Neuroscience and Basic Behavioral Science*) évoluent proportionnellement aux redistributions dans les différents autres domaines (comme celui de la recherche en épidémiologie), qui peuvent être défavorisés par la redistribution⁷¹. Autrement dit, si certains axes plus centrés sur les neurosciences ont un coût tel qu'ils engagent un risque de perte pour d'autres approches nécessaires à clinique et à la santé publique, ils pourraient être contre-productifs – car la santé de l'individu dépend toujours aussi d'une manière ou d'une autre de la santé de la population à laquelle il appartient (Giroux 2020). De ce fait, l'apport des neurosciences à la psychiatrie ne peut être valable et entendable que si un cadre général de travail tolérant peut être mis en place (Longino, 2006). Cette observation devrait rester cependant une ouverture qui dépasse le contexte épistémologique de notre thèse.

⁷¹ Même si les décideurs affirment proposer un portefeuille diversifié au sein de la recherche fondamentale, de la clinique et de la recherche portant sur les services, il est évident en pratique que les choix de financement à grande échelle en psychiatrie sont encore largement guidés par les promesses de réussite des sciences fondamentales.

Conclusion

Comme l'écrivait Cuthbert en 2020, « les RDoC représentent un mélange de l'ancien et du nouveau » (Cuthbert, 2020). Le projet RDoC réside ainsi dans le fait de « fournir des opportunités aux chercheurs face aux importantes questions que pose la recherche translationnelle fondée sur les neurosciences et les sciences du comportement » (Cuthbert, 2014). Par ses propositions, il ravive une problématique de longue date, cherchant à rapprocher les sciences du cerveau de la clinique psychiatrique (Hyman, 2007; Neumärker & Bartsch, 2003). Les RDoC sont loin de constituer un « changement de paradigme » ou une « révolution scientifique », comme cela a pu être avancé par un certain nombre de promoteurs de ce programme de recherche. Ils constituent un programme de recherche institué dans la continuité des programmes neuroscientifiques du XX^e et XXI^e siècle et qui présente une caractéristique commune avec le DSM : la volonté de structurer la psychiatrie autour d'éléments (neurobiologiques) justifiant l'existence de fonctions cognitives spécifiques (et de dysfonctions présumées dans le cadre du DSM). DSM comme RDoC cherchent à légitimer la psychiatrie en « opérationnalisant » (pour le DSM) ou en « fondant sur des données probantes issues de l'expérimentation et de l'épidémiologie » (pour le projet RDoC) les troubles psychiatriques (pour le DSM) ou les « fonctions cérébrales » (pour le projet RDoC). Le projet RDoC propose une matrice intégrant à la fois la psychologie et la psychopathologie sous forme d'éléments neuroscientifiques. Par ce mouvement, les RDoC rassemblent également des éléments épars de la psychiatrie au sein de dimensions désignant autant de variables latentes. Cette volonté de s'appuyer sur une telle approche par la dimension du projet RDoC est liée à une mise en tension que nous avons décrite au cours de ce chapitre : la volonté d'intégrer ce projet à la pratique clinique psychiatrique, tout en répondant aux multiples enjeux scientifiques (mais également sociaux, institutionnels et de financement du NIMH) qui ont été à l'origine de ce projet. Ainsi, en plus de concevoir le projet RDoC comme le produit d'une réponse palliative face à la déception d'autres projets précédents – témoin d'un « réalignement et d'un déplacement des intérêts de recherche » pour répondre aux échecs passés de la

psychiatrie (Rüppel, 2019) – nous proposons que ce projet soit le témoin de la mise en tension des enjeux cliniques et non-cliniques. En effet, nous avons vu que les notions d'intégration et d'interdisciplinarité, de comportement, de développement et d'environnement constituent des tentatives de rapprocher ce projet de recherche d'enjeux cliniques. Toute la seconde partie de ce chapitre met ainsi en avant la tension entre des fonctions psychologiques répondant à des enjeux de recherche et une volonté d'utilisation des RDoC en pratique clinique. Du fait de cette structuration fondée sur des dimensions, le projet cherche ainsi à résoudre cette tension, ce qui n'est réalisable qu'au prix de multiples adaptations et compromis temporels, organisationnels, réducteurs ou simplifiant. En utilisant la notion de fonction psychologique, conçue comme une dimension, les RDoC perpétuent l'idée qu'une variable latente devrait être absolument identifiée pour définir ce qu'est un trouble psychiatrique.

Alors que les différentes versions successives du DSM ont bien montré à quel point des initiatives bien intentionnées pouvaient « mal tourner » et devenir « embarrassantes » (Frances, 2009, 2013), dès lors que certaines imprécisions conceptuelles n'étaient pas résolues (rappelons la « confusion conceptuelle » de Wittgenstein – (Wakefield, 2014; Wittgenstein, 1973)), les RDoC semblent, curieusement, tomber de nouveau dans le piège d'une initiative nosographique amenant autant de difficultés qu'elle ne souhaite en résoudre. Dans le chapitre suivant, nous verrons que la HiTOP a cherché à s'ancrer sur des données probantes (comme le fait le projet RDoC), tout en conservant des dimensions constituées d'éléments cliniques (comme le fait le DSM constitué de signes et symptômes).

LA TAXONOMIE HIÉRARCHIQUE DE LA PSYCHOPATHOLOGIE :

À LA CROISÉE ENTRE CLINIQUE ET STATISTIQUES

Dans les deux chapitres précédents, en étudiant la tension entre les fonctions cliniques et non-cliniques du DSM et du projet RDoC, nous avons pu mettre en avant les défis à relever par l'approche par la catégorie / dimension. Rappelons que cette approche par la catégorie / dimension nécessite de concevoir une variable latente et fournit une définition d'une espèce psychiatrique en tant qu'espèce diagnostique (qui elle-même conçoit la définition de ce qu'est un trouble psychiatrique comme un ensemble de symptômes et d'autres spécificateurs).

Nous nous sommes appuyés sur le DSM pour illustrer l'importance du contexte historique dans lequel la discipline psychiatrique s'est « débattue » ces dernières décennies, afin de réduire cet écart entre impératifs cliniques et non-cliniques ; nous avons présenté le projet RDoC comme tentative de solution à cette disjonction clinique/non-clinique en mettant en avant ses efforts de cohérence entre neurosciences et pratique clinique. Malgré les limites et écueils du DSM relevés dès la parution de la troisième édition en 1980, l'approche fondée sur la catégorie / dimension (promulguant des espèces diagnostiques) a été essentielle pour promouvoir la recherche et l'épidémiologie, assurer la stabilité administrative, garantir les impératifs économiques et politiques. La discipline psychiatrique de cette dernière décennie semble donc reposer sur deux piliers : des diagnostics ordonnés au sein d'une classification utilisée plus ou moins explicitement en pratique clinique (le DSM) et une tentative scientifique de structurer la psychiatrie sur des bases neurobiologiques (le projet RDoC).

Cependant, pour ne citer que quelques-unes de leurs limites, le DSM présente des lacunes en termes de validité et peine à considérer la comorbidité entre les troubles ; le projet RDoC peine à être applicable en pratique clinique. Comme nous en avons fait l'hypothèse, la multiplicité de conceptions de la psychiatrie proviendrait d'une difficulté à répondre aux fonctions à la fois cliniques et non-cliniques requises par la discipline. C'est sur cet argument que nous soutenons, dans ce chapitre, que la HiTOP, publiée initialement en 2017, s'est présentée comme une classification alternative cherchant à se positionner à la croisée de ces fonctions. La HiTOP conçoit la psychopathologie comme un ensemble de dimensions cliniques organisées de manière hiérarchique, c'est-à-dire en continua « transdiagnostiques ». En même temps, elle se fonde sur une méthodologie statistique, celle de l'analyse factorielle, qui elle-même s'appuie sur la notion de variables latentes.

Ce chapitre décrit les structures de base de la HiTOP et son potentiel à s'intégrer à la pratique clinique et à la recherche psychiatrique. Nous illustrons les applications d'une telle approche dimensionnelle et hiérarchique par rapport à la nosologie traditionnelle. Nous discutons des obstacles à l'utilisation de la HiTOP dans l'environnement clinique et de recherche actuel en psychiatrie. Nous détaillons ensuite les potentiels apports de la HiTOP à la thérapeutique et à la compréhension mécanistique (neurobiologique) des troubles psychiatriques. Ces développements nous montrent finalement que la HiTOP, à l'image des conceptions et classifications décrites dans les chapitres précédents, constitue bien une approche par la dimension – fondée sur la notion de variable latente. Cet ancrage semble être au cœur des difficultés à réduire l'écart entre les fonctions cliniques et non-cliniques, défi sur lequel la HiTOP semble encore buter.

Introduction

En 2017, le consortium dédié à l'avancement de la taxonomie hiérarchique de la psychopathologie (*Hierarchical Taxonomy Of Psychopathology*, HiTOP) a proposé de développer une taxonomie fondée sur des données probantes, en mettant l'accent à la fois sur la fiabilité, la validité et la pertinence clinique (Kotov et al., 2017). L'article « fondateur » de la (HiTOP) a été publié en 2017 par Kotov et collaborateurs et a été amplement discutée dans deux issues spéciales de *World Psychiatry* en 2018 et en 2021 (Krueger et al., 2018; Lahey et al., 2021). Ses auteurs se réclament en partie du mouvement de la « nosologie quantitative », mouvement que nous décrirons par la suite. Nombre d'entre eux appartenaient au groupe de travail du DSM-5 sur les troubles de personnalité (2002 – 2013). D'ailleurs, en 2008, deux membres qui rejoindront plus tard le consortium HiTOP ont édité un livre important, intitulé « *Dimensional Approches in Diagnostic Classification* », dans le but de promouvoir la réflexion sur le programme de recherche dimensionnel du DSM-5 (Zachar et al., 2016). Comme son nom l'indique, le projet cherche à développer une « taxonomie » et non une classification figée, en considérant dans ce sens une taxonomie comme un projet plus large qu'une classification – à savoir une proposition initiale de hiérarchisation de dimensions, vouée à évoluer. Comme nous allons largement le détailler, la HiTOP est une classification issue de la validation de construits statistiques, qui sont à la fois dimensionnels et hiérarchiques. Elle est fondée sur le recueil de données collectées sur le terrain par des outils de psychométrie. En effet, le projet HiTOP décrit les troubles psychiatriques selon un continuum de variations quantitatives qui peuvent être organisées hiérarchiquement en plusieurs spectres et dimensions (Jablensky, 2018). La HiTOP propose une classification dimensionnelle hiérarchique verticale en six niveaux : super-spectre, spectres, sous facteurs, syndromes / troubles, traits et signes / symptômes, comme l'illustre la **Figure 5**.

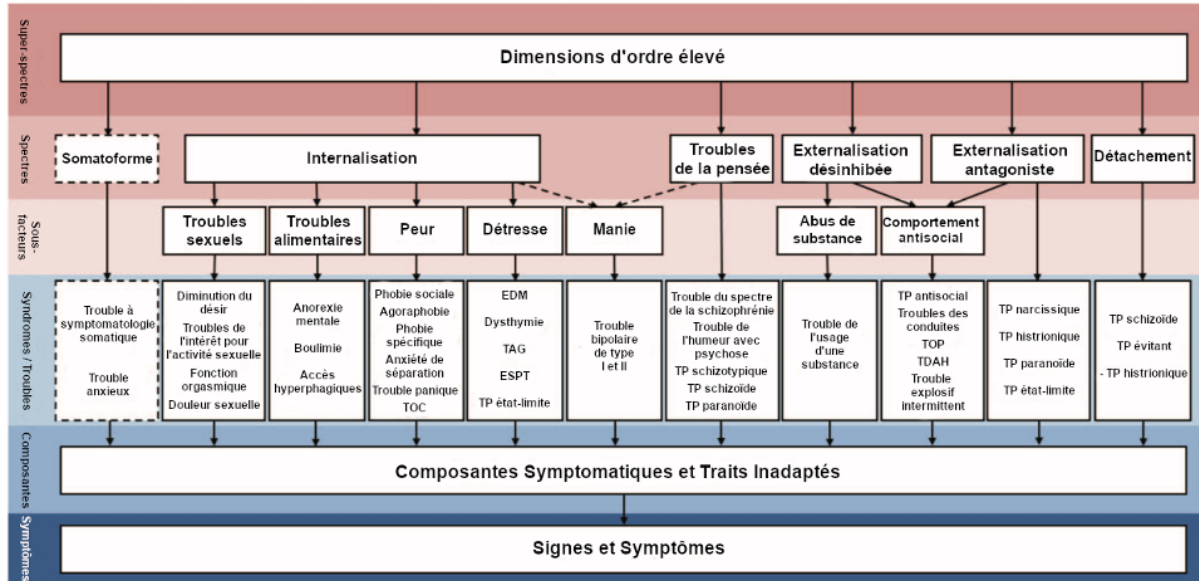


Figure 5 – Modèle du consortium de travail de la Taxonomie Hiérarchique de la Psychopathologie (*Hierarchical Taxonomy of Psychopathology*, HiTOP) (adapté de Kotov et collaborateurs). TAG – Trouble anxieux de séparation, TOC – Trouble obsessionnel-compulsif, EDM – Épisode dépressif majeur, TAG – Trouble anxieux généralisée, ESPT – État de stress post-traumatique, TP – trouble de la personnalité, TOP – Trouble oppositionnel avec provocation, TDAH – Trouble déficit de l’attention / hyperactivité, TEI – Trouble explosif intermittent.

Au-dessus du niveau le plus bas de la hiérarchie HiTOP constitué de signes et de symptômes, se trouvent des constellations de symptômes et des traits inadaptes coexistant fortement. Ces constellations et ces traits se regroupent eux-mêmes en syndromes. Les syndromes de la HiTOP ne correspondent pas nécessairement aux troubles du DSM-5 ou de la CIM-11 (World Health Organization, 2017), même s’ils constituent le niveau de la HiTOP qui est le plus similaire à ces troubles. Au niveau hiérarchique supérieur suivant se trouvent les sous-facteurs, reflétant de groupes de syndromes fortement liés statistiquement. Par exemple, un sous-facteur de « peur » comprend à la fois la phobie sociale, la phobie spécifique, l’agoraphobie et la pathologie évitante de la personnalité. Enfin, au-dessus de ce niveau hiérarchique, on retrouve des spectres, c’est-à-dire de larges groupes de sous-facteurs distincts les uns des autres, même s’ils restent interdépendants pour certains. Par exemple, la détresse, la peur, les sous-facteurs de l’alimentation et les sous-facteurs des troubles sexuels sont regroupés dans le « spectre d’internalisation » de la HiTOP. Au-delà, un facteur de psychopathologie générale (p) reflète une mauvaise adaptation

fonctionnelle globale (Caspi et al., 2014) et définit pour une part ce qui devrait relever de la psychopathologie d'un point de vue statistique. Ces caractéristiques structurelles de la HiTOP nous permettent déjà de comprendre en quoi la classification est une approche psychopathologique construite de manière hiérarchique.

Nous avons vu précédemment que lorsque le projet RDoC tente de s'extraire des difficultés qu'apportent le DSM, il persiste à considérer une variable latente en se structurant autour de dimensions (en l'occurrence, des fonctions psychologiques). Cependant, les difficultés d'une telle approche par la catégorie / dimension pourrait provenir de leur inclusion dans la matrice sur avis d'expert. Il en est de même pour le DSM, pour lequel les catégories ont été sélectionnées sur avis d'expert à la suite d'un travail qui a commencé au moins dans les années 1950 jusqu'à la publication du DSM-III en 1980. Or, au sein de la HiTOP, nous verrons que les dimensions de la classification sont issues de regroupements statistiques et non proposées sur avis d'experts. C'est donc une démarche plus empirique que théorique qui a permis de fournir de telles dimensions. Ces regroupements statistiques permettent-ils de mieux justifier le fait que des variables latentes (les dimensions) structurent la classification ? Le fait que les dimensions proviennent de calculs statistiques ne permettrait-il pas d'obtenir une classification répondant à des enjeux scientifiques, tout en répondant à un certain nombre de fonctions cliniques (du fait de l'ancrage de la HiTOP dans la sémiologie psychiatrique) ? Après avoir retracé l'ancrage historique de la HiTOP dans la nosologie quantitative et la psychologie dimensionnelle (I), nous explorerons les enjeux de validité et la capacité de la classification à répondre à des multiples fonctions cliniques et non-cliniques, à travers ces deux axes : le premier axe étudiera la structure dimensionnelle de la HiTOP (II) et le second axe analysera sa structure hiérarchique (III). Ces axes permettront de questionner la manière dont les approches dimensionnelles sont susceptibles d'améliorer la précision diagnostique, l'étude de la cooccurrence des troubles psychiatriques, l'élaboration d'un modèle déterminant un seuil pour les conditions pathologiques et la manière dont il est possible de relier des éléments physiopathologiques à des éléments cliniques. Dans un dernier temps, nous pointerons un certain nombre de limites inhérentes à la HiTOP, illustrant le fait

que, malgré sa méthode scientifique fondée sur les statistiques, la classifications tombe dans des difficultés semblables aux approches par la catégorie et dimension présentées précédemment, liées à l'utilisation de la notion de variable latente.

I. Psychologie dimensionnelle, nosologie quantitative et statistiques

1. La psychologie dimensionnelle

En psychiatrie, quatre ans après la parution de la troisième version du DSM-III, en 1984, Boyd et collaborateurs analysaient, dans une large cohorte de patients, qu'au sein du DSM existait une « tendance générale à la *cooccurrence*, de sorte que la présence de tout trouble augmentait les chances d'avoir presque n'importe quel autre trouble, même si le DSM-III ne le répertoriait pas comme un trouble connexe » (nous soulignons) (Boyd et al., 1984). Par exemple, les symptômes de panique seraient considérés comme une « caractéristique essentielle » du trouble panique, mais ils seraient également une « caractéristique associée » de la dépression caractérisée. Comme l'ajoutent les auteurs, « cette distinction linguistique, cependant, ne répond pas à la question de savoir pourquoi les symptômes de panique survenant au cours d'un épisode de dépression majeure devraient être considérés comme une manifestation de la dépression majeure et non comme une manifestation d'un véritable trouble panique survenant en même temps que la dépression majeure » (Boyd et al., 1984). Autrement dit, Boyd et collaborateurs montrent que le DSM-III établit, implicitement, une distinction entre les « caractéristiques essentielles » et les « caractéristiques associées » d'un trouble (Boyd et al., 1984). Comme nous allons le voir, des solutions à cette difficulté, inhérente aux classifications catégorielles, ont été proposées. C'est au travers de la dimensionnalité et de la hiérarchisation que la HiTOP a cherché à y répondre. Comme l'écrivait Fulford, de telles difficultés ne

sont pas nouvelles (Fulford & Handa, 2018). Elles ont été largement discutées dans les années 1960 et 1970. Kendell rappelait en 1975 que la psychopathologie pouvait difficilement être « découpée aux articulations » (Kendell, 1975), renvoyant à la possibilité de séparer des entités de manière naturelle (Markon et al., 2011).

Après Kraepelin et les premiers nosologistes, une grande partie de la recherche en psychopathologie a été cependant structurée autour des catégories et des dimensions. L'approche par la dimension est en partie issue de la psychologie dimensionnelle, que nous allons mieux détailler (Krueger et al., 2018). Les germes de la pensée dimensionnelle étaient perceptibles dans les classifications du XIX^e siècle. Par exemple, en 1860, Benoît-Augustin Morel conceptualisait les troubles psychiatriques comme des manifestations de la « dégénérescence », allant du « tempérament nerveux » à « l'idiotie » (Pichot, 1994). Dans son introduction, le DSM-5 proposait d'introduire une valence dimensionnelle : on y retrouve la double notion d'une dimension d'internalisation (représentant les troubles où prédominent les symptômes anxieux, dépressifs et somatiques) et d'une dimension d'externalisation (représentant les troubles où prédominent les symptômes impulsifs, de conduite disruptive et de dépendance à l'usage d'une substance). Cette ambition dimensionnelle adoptée au début de la constitution du DSM-5 n'a pas été poursuivie. Ce couple internationalisation / externalisation renvoie à deux des spectres de la HiTOP (spectre internalisé et spectre externalisé) et il est difficile de ne pas penser aux deux pôles psychanalytiques de la névrose et de la psychose – bien que cette analogie n'ait pas été testée de manière empirique et que la HiTOP se fonde avant tout sur des dimensions statistiques pour établir ces deux spectres (voir sections suivantes). La HiTOP hérite donc en partie des réflexions sur la dimensionnalité qui ont été élaborées en psychologie au cours du siècle précédent⁷².

⁷² Plus généralement, l'approche dimensionnelle a fait l'objet d'une abondante littérature en psychiatrie et en philosophie de la psychiatrie, qui ne sera pas discutée ici (voir par exemple, en français, Demazeux, 2008).

2. La nosologie quantitative

Les premières tentatives d'étudier les classifications de manière mathématique (statistique) et hiérarchique ont commencé à se développer à la fin du XIX^e siècle avec Peano et Cantor, puis avec Peirce (Peirce, 1880) et Schröder (Schröder, 1890). Au milieu du XX^e siècle, des modèles mathématiques de classifications hiérarchiques étudiés aux États-Unis par Sneath et Sokal (Sneath & Sokal, 1973) ou en Angleterre par Jardine et Sibson (Jardine et al., 1971) ont été développés en France, avec les travaux de Barbut et Monjardet (Barbut & Monjardet, 1970) ou Benzécri (Benzecri, 1975). Ces modèles proposaient d'établir une classification empirique en évaluant les similitudes ou les dissemblances entre les éléments à classer, en se fondant sur un modèle de la taxonomie tel qu'il avait par exemple été décrit par Buffon dès 1749 (Buffon, 1749). La construction de classifications empiriques statistiques et hiérarchiques, évaluant les ressemblances et différences entre catégories et dimensions, a progressivement imprégné différentes disciplines scientifiques, dont la médecine (Hempel, 1955). Depuis au moins cinq décennies, l'intégration des modèles statistiques et hiérarchiques s'est concrétisée autour de la « nosologie quantitative », principalement soutenue par des chercheurs en psychologie s'intéressant aux troubles de personnalité. Ce mode de pensée cherche à identifier les constellations de signes, symptômes et comportements pour structurer la médecine et la psychopathologie (Blashfield, 1982; Eysenck, 1944; Lorr et al., 1963; Wittenborn, 1951). Cette identification passe par des analyses statistiques permettant le regroupement de signes, symptômes et comportements en composantes et traits dimensionnels, eux-mêmes regroupés en dimensions plus larges, de manière hiérarchique. Par exemple, les deux spectres d'intériorisation et d'externalisation ont été identifiés bien en amont de la constitution de la HiTOP, au sein d'une série d'articles publiés dans le champ de la nosologie quantitative (Achenbach et al., 2017). D'ailleurs, un troisième spectre, le trouble de la pensée, avait également été identifié de manière statistique (Eaton et al., 2013; Klein et al., 2011; Markon, 2010). Les travaux de la nosologie quantitative en psychiatrie s'appuient notamment sur les analyses factorielles, retrouvée par exemple chez Eysenck dès

1944 (Eysenck, 1944) ou Achenbach dès 1966 (Achenbach, 1966). Ces travaux fournissent une analyse quantitative de la cooccurrence des symptômes. En statistiques, le fait de déterminer une dimension sur la base de la concurrence de symptômes s'appelle justement une variable latente, défini comme une entité non observable dans les données. Cette cooccurrence permet d'établir des dimensions latentes censées refléter la « structure des problèmes psychologiques » (Achenbach, 1966). Puisque la nosologie quantitative s'appuie sur la psychométrie, il faut discuter de la nature des signes ou des symptômes recueillis. Il est courant de penser la définition du symptôme en psychopathologie comme étant fortement dépendante de la conception (classification) dans laquelle il s'insère (Berrios, 1996). Nous pourrions parler de dépendance de la sémiologie à la nosologie. Les construits proposés par la HiTOP font nécessairement émerger une conception du symptôme différente de celle qui a été amenée au fil des révisions du DSM⁷³. La HiTOP n'a pas cherché à redéfinir les signes et symptômes de la psychopathologie. Ce conservatisme, prudent, a été adopté pour au moins trois raisons : premièrement, le fait de se fonder sur le répertoire psychiatrique déjà isolé depuis plusieurs décennies par la psychiatrie moderne et contemporaine faciliterait la fiabilité, la communication entre chercheurs et la compréhensibilité de la classification ; deuxièmement, les symptômes évolueront en fonction de l'utilisation pratique des chercheurs et cliniciens ; troisièmement, ils changeront en fonction de la dépendance de la sémiologie à la nosologie, c'est-à-dire en fonction de la structure même de la HiTOP.

3. Classification numérique statistique

Dans le DSM, le regroupement de symptômes au sein d'une catégorie est conçu comme le résultat d'une entité non observable, hypothétique. Cette entité hypothétique n'a été validée qu'*a posteriori*, par des

⁷³ De plus, les symptômes évoluent en fonction des valeurs normatives d'une société et de l'ancrage culturel des communautés de chercheurs et des cliniciens qui utilisent une classification (Jablensky, 1999).

études de fiabilité. La HiTOP est une classification numérique (Sokal et Sneath, 1963), c'est-à-dire qu'elle opère des regroupements par des méthodes numériques en fonction des caractères observables (Hull, 1988). Les méthodes statistiques utilisées lors du processus de construction de la HiTOP correspondent à des analyses en classes latentes et notamment des méthodes d'analyse factorielle. Le principe de l'analyse factorielle est de supposer qu'un certain nombre fini de variables observées (ici, les signes et les symptômes) possèdent des schémas de réponses similaires du fait de leur regroupement au sein d'une variable latente. Ce regroupement s'effectue par la projection des variables dans un espace de dimension inférieur au nombre de variables, en utilisant un degré de similarité entre variables. Parmi ces analyses factorielles, on trouve par exemple les analyses en composantes principales, les composantes principales étant des variables latentes⁷⁴. **Nous avons proposé de définir une variable latente comme une entité qui ne peut être déduite qu'indirectement à partir d'autres variables observables ou mesurables.** Il s'agit donc d'un « construit » statistique non directement mesuré et non contenu dans l'ensemble de données. Autrement dit, la variable latente est une entité *hypothétique*, construite, censée sous-tendre l'apparition et/ou le maintien des symptômes cliniques observables. Dans le cadre de la HiTOP, il s'agit d'une dimension (factorielle), dont la dénomination ne dépend que de l'ensemble des parties que ce construit rassemble. Dans d'autres contextes, la variable latente peut correspondre à une *cause commune*, par exemple lorsqu'elle rassemble un ensemble d'éléments biologiques : la variable latente correspond alors au *mécanisme* qui rassemble ces éléments. De fait, sur le plan ontologique, une variable latente est moins engagée qu'une cause commune, c'est-à-dire qu'elle nécessite moins que la

⁷⁴ De manière intéressante, la théorie qui sous-tend la structure des questionnaires de psychométrie correspond à la « théorie de réponse aux items » (*Item Response Theory*, [IRT]) (Embretson & Reise, 2013), développée dans la deuxième moitié du XX^e siècle. Cette théorie permet de hiérarchiser un ensemble de questions (les items des questionnaires) en s'affranchissant de l'influence des patients de référence. Pour cela, elle regroupe les items en postulant l'existence d'une variable latente. Ainsi, dans la structure même des questionnaires qui ont permis de constituer la HiTOP, des variables latentes sont utilisées. Le processus de construction de la HiTOP a également nécessité de déceler les variables latentes utile à la structuration de la classification (puisque le principe même des analyses factorielles est de trouver un certain nombre de variables latentes expliquant la variance du jeu de données).

cause commune de postuler certains attributs, une essence ou un univers de discours donné. Elle représente plutôt le résultat d'une mesure de distance entre symptômes.

Pour résumer, l'analyse factorielle permet donc de déterminer et de hiérarchiser des variables latentes (non-observables) qui résument les données cliniques (observables). Les dimensions de la HiTOP correspondent à ces variables latentes. La HiTOP est donc fondamentalement ancrée sur un postulat statistique possédant un corrélat conceptuel : des variables latentes dimensionnelles, correspondant à des espèces diagnostiques, sont nécessaires et / ou utiles pour regrouper des signes et des symptômes. À l'instar des deux autres conceptions de la psychiatrie présentées précédemment, la HiTOP s'apparente à une approche par la catégorie / dimension. Dans la deuxième section, nous analyserons plus précisément la dimensionnalité qui est à l'œuvre au sein de la HiTOP.

II. La structure dimensionnelle de la HiTOP

1. Description de la dimensionnalité

La HiTOP propose de concevoir les troubles en termes de dimensions plutôt qu'en termes de catégories. Elle est construite par regroupement des symptômes les plus élémentaires, permettant de délimiter des niveaux de généralité de plus haut degré (par exemple, des sous-facteurs ou des spectres) jusqu'à englober, au plus haut degré de la hiérarchie, tous les troubles psychiatriques. Par exemple, le spectre d'internalisation englobe des sous-facteurs comme la peur et la détresse qui, à leur tour, contiennent des phénotypes cliniques plus étroits encore, comme l'anxiété sociale ou l'insomnie (voir **Figure 5**). Se déclarant dimensionnelle, la HiTOP affirme traiter les phénomènes psychiatriques de manière « continue ». Cela signifierait qu'en étant intégré au sein d'un super-spectre, d'un spectre, d'un sous-

facteur, d'un syndrome ou d'une composante, chaque patient pourrait évoluer selon un continuum pathologique. Cependant, ce qui est réellement dimensionnel dans la HiTOP reste ambigu. Deux conceptions de la dimensionnalité semblent se dissimuler derrière cette acception. Il pourrait s'agir en premier lieu d'un glissement d'usage à partir du terme de « dimension » factorielle utilisé en statistiques. En effet, comme décrit plus haut, les analyses factorielles réduisent le nombre de variables en les rassemblant au sein d'un nombre plus restreint de dimensions⁷⁵. Au sein de chacune de ces dimensions, il peut y avoir ou non tel ou tel signe et symptôme. Par exemple, une dimension intégrera plutôt les symptômes de bradypsychie, d'aboulie et d'anergie alors qu'une autre dimension intégrera plutôt les symptômes de craving, de sevrage ou de tolérance. Un patient sera ainsi classé dans une de ces deux dimensions, même s'il ne présente pas tous les symptômes. On parle donc d'approche dimensionnelle parce qu'en considérant un certain nombre de patients appartenant à la même dimension, tous n'auront pas les mêmes symptômes ; certains en auront moins, ils seront donc « dimensionnellement » considérés moins sévères, d'autres en auront plus, ils seront donc considérés plus sévères. Si un patient est intégré à un sous-facteur, il appartient à une dimension contenant moins de symptômes que s'il avait été intégré à un spectre⁷⁶. Il s'agit donc d'une forme de dimensionnalité qui concerne les construits de la HiTOP, par opposition à une forme de dimensionnalité concernant l'intensité des symptômes – que nous allons détailler ci-dessous.

Nous soutenons que la dimensionnalité de la HiTOP se situerait au niveau des symptômes, qui peuvent être eux-mêmes exprimés plus ou moins intensément. Les échelles de psychométrie considèrent en effet un continuum symptomatique, en cotant par exemple un symptôme (un item) de 1 à 5 sur une échelle de

⁷⁵ Sur le plan statistique, chaque dimension correspond à un plan sur lequel les données sont projetées et chacun de ces plans a pour objectif de maximiser la variance des données (pour ne perdre que le minimum d'information lors de cette projection). Ainsi, les « dimensions » au sens statistique n'ont pas la même valeur (épistémique) que les dimensions conceptuelles utilisées dans la nosologie psychiatrique.

⁷⁶ Cette observation pourrait être nuancée car tout patient intégré à un sous-facteur est conceptuellement et nécessairement intégré au spectre supérieur. Cependant, ici, nous nous intéressons à l'utilisation de la HiTOP en pratique clinique, lors de laquelle une seule dimension est utilisée à la fois pour un patient. En d'autres termes, philosophiquement, l'espèce psychiatrique définie par la HiTOP pourrait être très large (par exemple, l'ensemble des conditions contenues sous le facteur p de la psychopathologie) ou très restreinte (par exemple, les syndromes ressemblant le plus aux catégories du DSM).

Likert⁷⁷. Cette conception est bien retrouvée dans les descriptions des diagnostics dimensionnels, par exemple lorsqu'il est fait mention qu'« un diagnostic dimensionnel a au moins trois valeurs ordinales, qui peuvent aller d'une échelle à trois degrés (au minimum) jusqu'à une échelle continue » (Lopez et al., 2007). Par exemple, l'anxiété sociale est un continuum allant des interactions sociales « confortables » à la détresse sociale survenant dans un grand nombre de contextes. Dans une telle classification dimensionnelle, il ne s'agit donc pas de considérer la présence ou l'absence d'un symptôme dans une dimension, mais plutôt de considérer l'intensité des symptômes qui la constitue. Les dimensions intègrent des symptômes exprimés à des intensités différentes en fonction des patients, plutôt que des ensembles différents de symptômes. En neuropsychologie, le quotient intellectuel est évalué de cette manière : à cinq construits (les cinq scores du test de quotient intellectuel – QI) est associée une échelle numérique (les points des cinq scores). Un QI inférieur à 2 écarts types par rapport à la moyenne (c'est-à-dire dans la fourchette 85 – 70) peut permettre un diagnostic et ainsi ouvrir des droits à l'allocation de ressources adapté au niveau de sévérité spécifié (American Psychiatric Association, 2013). Dans l'article de 2017 du consortium de la HiTOP, Kotov et collaborateurs (Kotov et al., 2017) soulignaient que « l'imposition d'une nomenclature catégorielle à des phénomènes naturellement dimensionnels entraîne une perte d'informations importante et une instabilité diagnostique » et ils poursuivaient : « (l) a prise de conscience grandissante de la présence de phénomènes psychotiques à certains degrés dans la population générale ainsi que dans un large éventail de patients en psychiatrie ambulatoire va dans le sens d'une répartition de ces symptômes dans la population générale qui ne se résume pas à une pathologie sévère et qualitativement distincte » (Kotov et al., 2017). Avec les notions de « degré » et de « répartition [égale] des symptômes », on retrouve l'idée d'une dimensionnalité qui devrait être décrite au niveau des symptômes. Nous verrons dans la seconde partie de cette thèse en quoi ce souci pour la description de l'intensité des symptômes est d'un grand intérêt pour mieux comprendre la psychopathologie. Pour conclure, dans la HiTOP, du fait de

⁷⁷ Certes, les catégories du DSM acceptent aussi d'intégrer un même symptôme exprimé plus ou moins intensément chez un patient. Mais la spécificité de la HiTOP est que celle-ci serait intrinsèquement construite sur ces différences d'intensité.

l'évolution continue des signes et symptômes, il n'est plus (seulement) question d'étudier la présence ou l'absence d'un symptôme (comme « absence de tristesse » versus « tristesse intense »), mais d'étudier un symptôme le long d'un continuum de plus ou moins faible intensité.

Mais que fait exactement le clinicien qui recueille des symptômes pour utiliser la HiTOP ? Un certain nombre d'études tendent à montrer que les approches dimensionnelles permettraient de mieux aligner la nosologie avec les données cliniques qui organisent la psychopathologie et que recueillent le clinicien (Krueger & Bezdjian, 2009). En quelque sorte, la pratique clinique s'apparente à une « analyse factorielle » heuristique du soin : le soignant récupère des données pour les « placer » dans une dimension ou une catégorie préalablement définie. Certes, dans cette conception (dimensionnelle) du soin, il n'est pas question de rechercher une « espèce naturelle » sous-jacente aux ensembles de symptômes. Cependant, une telle approche dimensionnelle nécessite toujours de présupposer une variable latente, conçue comme une entité non-observable qui permet de structurer les symptômes recueillis. Malgré sa dimensionnalité, la HiTOP constitue bien une approche par la dimension/catégorie.

III. La structure hiérarchique de la HiTOP

1. Description de la hiérarchisation

Dans le cadre de la publication des critères de Feighner en 1972 (Feighner et al., 1972) et des critères diagnostiques pour la recherche (Spitzer et al., 1978), qui ont soutenu la constitution du DSM-III en 1980, il s'agissait de rejeter le concept de hiérarchie diagnostique au motif qu'en l'absence d'étiologie, aucune hiérarchie n'était envisageable : sans connaissance des causes, il ne serait pas possible de dire quel niveau devrait être préféré par rapport à un autre. Mais cette perspective ignore un certain nombre de contraintes

de la pratique clinique. Par exemple, comme décrit précédemment au travers du texte de Boyd (1984), l'absence de hiérarchisation ne permet pas de prendre en compte l'importance du nombre de symptômes qui doit être présent pour diagnostiquer un trouble. Si un symptôme survient plus fréquemment dans un trouble que dans un autre, le symptôme est plus spécifique du premier que du deuxième. Le DSM ne fournit pas d'outils pour mesurer cette spécificité (Ghaemi, 2018). Au contraire, du fait de sa structure basée sur des regroupements plus ou moins importants de signes et de symptômes, la HiTOP peut se passer de l'étiologie pour proposer une hiérarchisation. Elle est une classification hiérarchique parce qu'elle regroupe les dimensions entre elles et non pas du fait de l'appartenance de ces dernières à une hypothétique étiologie commune.

La hiérarchie de la HiTOP est verticale : les signes et les symptômes sont « en bas » de la classification et le facteur p (le facteur désignant ce qui relève de la psychopathologie) est « en haut » (voir **Figure 5**), ce qui veut dire que plus on « remonte » dans la classification, plus les dimensions sont larges. Plus les dimensions sont larges, plus elles sont riches en termes de quantités de symptômes. Par exemple, un syndrome peut contenir neuf symptômes, tandis qu'un spectre, plus haut dans la hiérarchie, contiendra tous les symptômes de tous les syndromes qu'il réunit (par exemple deux fois neuf symptômes). La hiérarchisation de la HiTOP impose que chaque niveau inférieur soit contenu dans le niveau supérieur. La hiérarchisation au sein de la HiTOP est donc dite « ascendante » (*bottom-up*). Dans le processus de construction de celle-ci, les symptômes d'un certain nombre de patients soumis à des échelles psychométriques ont été recueillis et regroupés en dimensions de plus en plus larges. Il existe donc un processus de découpage (*split*) plus ou moins sévère à chaque niveau⁷⁸. Cette observation est importante pour définir une espèce psychiatrique : celle-ci peut être entendue comme une dimension dont le nombre de symptômes varie en fonction du choix pragmatique du niveau de hiérarchisation.

⁷⁸ Au sein du DSM, le regroupement de signes et de symptômes dans une même catégorie s'est fait sur l'avis des cliniciens au fil de l'histoire de la psychiatrie. Il s'agit alors d'une approche dite « descendante » (*top-down*), qui s'est construite sur la base de l'histoire de la discipline. L'étude d'appariement des catégories du DSM sur des échantillons cliniques a été réalisée dans un second temps, au sein des différentes vagues d'études sur le terrain (*field-trials*).

2. Hiérarchisation, clinique et comorbidité

La hiérarchisation permettrait une meilleure prise en compte de la comorbidité. Pour un patient donné, elle offrirait la possibilité de choisir une dimension qui regroupe convenablement les symptômes, sans faire nécessairement appel à des entités comorbides comme le fait le DSM. Par exemple, un trouble panique ne serait pas « comorbide » à un épisode dépressif caractérisé, mais une dimension supérieure, qui intégrerait ces deux troubles, serait utilisée en pratique clinique. En effet, la HiTOP rassemblerait à un niveau supérieur des symptômes qui ne seraient pas liés au sein de niveaux inférieurs. De cette manière, elle ferait de la comorbidité une de ses caractéristiques explicites en classant dans une dimension supérieure (par exemple, le spectre d'internalisation) des dimensions considérées comme comorbides par le DSM (par exemple, le trouble panique et l'épisode dépressif caractérisé, renvoyant à une « comorbidité artificielle », une expression de Michael First pour désigner une des conséquences imposées par le DSM qui divise les diagnostics catégoriels (First, 2005)). Les symptômes étroitement corrélés seraient affectés à la même dimension (par exemple, le spectre d'externalisation désinhibée) tandis que les symptômes non fortement corrélés seraient affectés à différentes dimensions (par exemple, le sous-facteur usage de substance et le sous-facteur comportement antisocial). Ainsi, au sein de la HiTOP, les comorbidités rencontrées en clinique pourraient contribuer à mieux expliquer la psychiatrie plutôt qu'elles ne l'affaibliraient (van Loo et al., 2013). Autrement dit, la HiTOP accueillerait mieux la comorbidité. La hiérarchisation permettrait également d'affiner la « granularité » de l'analyse clinique. En effet, en fonction de la question posée et de l'objectif désiré, le clinicien pourrait s'appuyer sur un spectre très large ou sur un sous-facteur moins important. La hiérarchisation de la HiTOP permettrait d'analyser plus finement les troubles (notamment tels qu'ils sont décrits au sein du DSM), en explorant leurs composants isolément (Simms et al., 2008). Par exemple, le trouble panique pourrait être décomposé en symptômes

physiologiques (par exemple, tachycardie et sensations d'étouffement) et psychologiques (par exemple, pensées de mort). La dimension inférieure de la classification (voir **Figure 5**) permet de ne prendre en considération qu'une dimension « symptomatique ».

Ensuite, la HiTOP permettrait d'évaluer le pouvoir prédictif de plusieurs dimensions étudiées ensemble (Waszczuk et al., 2017). Plutôt que d'utiliser des protocoles distincts pour traiter des diagnostics indépendants, tels que l'épisode dépressif caractérisé et l'état de stress post-traumatique, la HiTOP pourrait permettre d'étudier un spectre large, fournissant un protocole unifié, transdiagnostique par rapport aux catégories du DSM (Barlow et al., 2017; Waszczuk, 2021).

Enfin, en termes thérapeutiques et d'orientation dans les soins, la nature hiérarchique de la HiTOP permet aux cliniciens d'adopter une approche plus progressive. Par exemple, en recevant un patient, un clinicien pourrait d'abord utiliser un des niveaux les plus élevés (notamment le facteur p de la psychopathologie, situé au niveau du super-spectre) pour affirmer qu'il s'agit d'un patient relevant de la psychiatrie, puis il affinerait son diagnostic et sa prise en charge « en cascade » vers le bas (en fonction du temps, des moyens et des besoins des cliniciens ou des services de soin) (Lahey et al., 2021). C'est déjà ce qui est réalisé implicitement dans le parcours de soin classique d'un patient consultant dans une institution psychiatrique. Par exemple, l'aspect pathologique (facteur p de la psychopathologie) est endossé aux urgences en analysant le préjudice décrit par le patient, sa souffrance, sa détresse ou son handicap; dans la suite du parcours de soin du patient, en hospitalisation, un diagnostic est posé (par exemple, au niveau des spectres); après la sortie d'hospitalisation, il pourra consulter un centre expert où un syndrome plus affiné sera retenu. Du fait de cette hiérarchie, l'utilisation de la HiTOP est donc relativement souple en fonction des contraintes cliniques pragmatiques.

Dimensionnalité et hiérarchisation sont intimement reliées par un enjeu décisionnel diagnostique et interventionnel : lorsqu'on atteint les dimensions supérieures de la classification, les dimensions diminuent en granularité diagnostique (correspondant à la capacité à désigner un patient par une dimension précise et restreinte), tout en augmentant la quantité de symptômes qu'elles regroupent (il y a plus de

symptômes contenus dans une dimension supérieure que dans une dimension inférieure). Pour résumer, la HiTOP fournit des modèles statistiques et hiérarchiques qui permettent une analyse quantitative de la cooccurrence des symptômes. Elle permet d'affiner la granularité de l'analyse clinique par l'exploration concomitante de plusieurs dimensions, en adaptant le niveau hiérarchique en fonction des contextes cliniques. Dans la section suivante, nous allons explorer les enjeux et défis liés à ces différentes caractéristiques.

IV. Discussion des enjeux

À l'avenir, l'approche de la HiTOP devra relever plusieurs défis. Ceux-ci relèvent de quatre principaux enjeux : un enjeu de validité, un enjeu méthodologique, un enjeu inhérent à la dimensionnalité et un enjeu d'utilité clinique.

1. Enjeux de validité

Au fil de ce chapitre, nous comprenons que la HiTOP a été structurée en réponse aux limites et écueils du DSM et du projet RDoC. Il existe une déception nosologique liée à cinquante ans d'expérience avec des catégories psychiatriques descriptives (du DSM), fournissant peu de perspectives étiologiques. Cette déception est en partie liée au fait que, contrairement aux affirmations faites dans les années 1970 et 1980, la fiabilité du diagnostic (qui permet la communication entre cliniciens et la reproductibilité dans le temps) n'a pas conduit à améliorer la validité du diagnostic. En effet, dans le DSM, la fiabilité est devenue une fin en soi, conduisant d'ailleurs à une certaine réification des diagnostics psychiatriques (Ghaemi, 2018). Le DSM répond à des fonctions cliniques – même si ce n'était pas son objectif initial – du fait de

l'acceptable fiabilité de la plupart de ses catégories ; cependant, le DSM n'est pas fondé sur des données extraites du terrain. De son côté, le projet RDoC s'appuie bien sur des données mises en évidence de manière empirique ; cependant, il reste peu applicable en pratique clinique. La HiTOP semble chercher un équilibre entre fiabilité et validité en tentant de s'appuyer sur les forces de chacune de ces précédentes conceptions. Les promoteurs de la HiTOP avancent en effet que les phénotypes dimensionnels possèderaient une fiabilité importante, tout en étant construits sur la base d'un recueil psychométrique, censé assurer une certaine validité.

La validité pourrait également être soutenue par un alignement des dimensions de la HiTOP avec différents mécanismes étiologiques des troubles (Andrews et al., 2009; Keyes et al., 2012). Nous allons détailler ce point, qui renvoie pour une part aux enjeux de validité et nous intéresse donc pour notre question principale. En effet, si du fait de sa structure (dimensionnelle et hiérarchique), la HiTOP parvenait à une validation satisfaisante en mettant en relation ses dimensions avec des mécanismes biologiques, l'approche par la dimensions/catégorie pourrait retrouver une certaine envergure ou stature dans le champ de la nosologie psychiatrique. Elle permettrait de répondre à des fonctions non-cliniques (de recherche notamment) tout en répondant à des fonctions cliniques, du fait de son approche dimensionnelle.

Les apports de la dimensionnalité et de la hiérarchisation de la HiTOP peuvent soutenir l'identification de biomarqueurs (Lahey et al., 2021). La mise en évidence de biomarqueurs mécanistiques en psychiatrie est restée très modeste ces dernières décennies. La difficulté à mettre en évidence des biomarqueurs tiendrait pour une part aux lacunes des approches diagnostiques catégorielles plutôt qu'aux limitations intrinsèques des approches biologiques (Totah et al., 2016). En effet, les diagnostics catégoriels ne seraient pas optimaux pour identifier des mécanismes psychopathologiques, notamment du fait d'une prise en compte peu pertinente de la comorbidité, de l'hétérogénéité intra-trouble (deux patients peuvent recevoir le même diagnostic même s'ils ne partagent presque aucun symptômes), de leur modeste fiabilité et, entre autres, des seuils utilisés pour définir les diagnostics – qui manquent de justification empirique (Stephan et al., 2016). Certes, comme nous l'avons vu dans le chapitre précédent, des cadres agnostiques comme les RDoC

(Cuthbert, 2014) ou la *NIAAA Addictions Neuroclinical Assessment* (Kwako et al., 2017) ont tenté de répondre à ces insuffisances. Cependant, leur relation avec la clinique et le niveau de fonctionnement en vie quotidienne reste modeste. La revendication de recherche de mécanismes causaux sous-jacents aux dimensions de la HiTOP est née de la faiblesse évidente du DSM à trouver une unique voie de causalité associée à une catégorie. La hiérarchisation permettrait de corrélérer des dimensions de niveau supérieur (par exemple, les spectres d'internationalisation et d'externalisation) à des marqueurs qui ne peuvent pas correspondre à des dimensions de niveau inférieur, contenant moins de composants. De même, l'aspect dimensionnel des symptômes évoluant par degrés le long d'un continuum permettrait de mieux corrélérer des marqueurs continus aux dimensions cliniques continues. L'utilisation de dimensions de niveau supérieur favoriserait la découverte de mesures neurobiologiques étiopathogéniques. En couvrant plusieurs diagnostics, les dimensions de niveau supérieur de la HiTOP soutiendraient la recherche de mécanismes reliés aux dimensions « comorbides ». Pour fournir un exemple de ce qu'apporte la HiTOP aux enjeux de recherche des biomarqueurs et de validité, sur le plan thérapeutique, les inhibiteurs sélectifs de la recapture de la sérotonine se sont révélés efficaces pour différentes conditions d'internalisation (Martinez et al., 2008) et pour plusieurs sous-facteurs du spectre (notamment l'anxiété, la dépression ou les troubles alimentaires). Cependant, il ne faudrait pas que ces preuves d'efficacité constituent un argument pour la validation des spectres – raccourci qui a pu être préjudiciable au DSM pour certaines de ces catégories (Horwitz & Wakefield, 2007). D'autres spectres peuvent également être ciblés par des techniques ayant des effets larges, par exemple les entretiens motivationnels pour des spectres extériorisés désinhibés (Lundahl et al., 2010). Une étude en imagerie de perfusion cérébrale⁷⁹ chez plus de 1 000 patients a révélé que le facteur général de la psychopathologie (le facteur p de la HiTOP) était associé à une perfusion élevée dans la cingulaire antérieure (Latzman & DeYoung, 2020; Waszczuk, 2021). Plus précisément, si un tel marqueur biologique est susceptible de conférer un risque à un niveau plus élevé

⁷⁹ L'imagerie de perfusion cérébrale permet de distinguer le tissu cérébral hypoperfusé du tissu définitivement ischémique, dans le cadre des accidents vasculaires cérébraux en phase aigüe.

(celui du facteur p), il influencera un plus large éventail de comportements que des marqueurs reliés à des niveaux inférieurs. Cependant, même après des décennies de recherche et une littérature abondante, les relations entre les dimensions de personnalité et des mécanismes physiopathologiques sous-jacents sont encore difficiles à identifier. La recherche sur les dimensions de personnalité a en effet permis d'identifier de nombreux ensembles de marqueurs, fournissant une validité en termes mécanistiques à celles-ci (Kendler & Myers, 2010; Skodol et al., 2011). Mais leur application en pratique clinique reste faible, voire inexistante. Il se pourrait que la HiTOP, malgré une identification précise des corrélats entre dimensions et biomarqueurs, rencontre des limites similaires à celles de la recherche sur les dimensions de personnalité. La validité des dimensions ne dépend certes pas de leur utilité clinique. Mais il se peut que la HiTOP ne fasse seulement pas mieux que le DSM, en ne faisant que vérifier que ses dimensions soient plus ou moins reliées à des biomarqueurs fort peu utiles pour la pratique clinique. Cela, sans compter que cette association entre dimensions et étiopathogénie est encore hautement programmatique.

2. Enjeux méthodologiques

Nous l'avons vu plus haut, la construction de la HiTOP s'est appuyée sur un recueil des données psychométriques (lui-même dépendant des théories sous-jacentes à la construction des échelles psychométriques comme la théorie de réponse aux items)⁸⁰ et sur l'interprétation qualitative des dimensions extraites des analyses statistiques (Wittchen & Beesdo-Baum, 2018).

Or ces échelles comportent de nombreuses limites, tant en termes de validité que de fiabilité ou de pertinence clinique. Comme le soulignaient Wittchen et Beesdo-Bau en 2018, la HiTOP est fondée presque exclusivement sur des mesures issues d'études transversales. Ces mesures illustrent « une image

⁸⁰ La théorie de réponse aux items est un modèle statistique qui permet d'évaluer les propriétés des items d'un questionnaire, par exemple en analysant la difficulté ou la discrimination avec laquelle ils sont remplis par les patients – une méthode statistique qui se fonde sur une variable latente afin de comparer les items entre eux.

instantanée subjective et verbale du niveau de souffrance symptomatique qu'une personne signale au moment de l'enquête » (Wittchen & Beesdo-Baum, 2018)⁸¹. Pour Zachar, un des inconvénients du « discours révolutionnaire » utilisé pour présenter la HiTOP (Krueger et al., 2018) tient au fait qu'il met l'accent sur la discontinuité entre le passé et le présent, détournant l'attention de nombreuses continuités, par exemple de certaines limites qui ont toujours été floues entre le modèle médical psychiatrique et le modèle psychologique⁸² (Zachar, 2018).

Jablensky souligne une limite de la HiTOP plus profonde encore : l'analyse et l'interprétation purement statistiques des données ne remplacent pas une étude comparative bien conçue des populations cliniques évaluées (qui incluent le choix du traitement, le pronostic et l'état fonctionnel des participants) (Jablensky, 2018). Autrement dit, la HiTOP possède des dimensions appuyées par les statistiques, mais celles-ci ne sont pas nécessairement pertinentes pour la pratique clinique. En effet, rien ne permet d'affirmer que la délimitation de dimensions sur la base de méthodes statistiques soit utile pour la pratique clinique. Extraire des dimensions par le regroupement d'items pourrait mettre de côté certains symptômes importants (par exemple, les idéations suicidaires pourraient n'être que rarement retrouvées, alors qu'il s'agit d'un symptôme particulièrement important à savoir rechercher dans toutes les dimensions) ou au contraire agréger des symptômes sans que cela n'ait d'utilité clinique (par exemple, l'hypoprosexie dans les troubles du neurodéveloppement n'est véritablement à rechercher que dans certains phénotypes en fonction du bilan neuropsychologique).

⁸¹ De plus, comme le font remarquer ces auteurs, aussi attrayante et impressionnante que puisse être la représentation visuelle d'une nouvelle taxonomie de la psychopathologie, l'utilisation de « nouveaux mots pour d'anciens pourrait augmenter le risque que les résultats de recherche déjà établis ne soient pas pris en compte à l'avenir ». En effet, un certain nombre de termes nouveaux sont introduits dans la HiTOP, par exemple la notion de détachement ou d'externalisation désinhibée, qui nécessitent une appropriation par le clinicien.

⁸² Nous avons souhaité ne pas entrer dans les débats sur les définitions et les frontières de ces modèles, qui relèvent plutôt de l'étude (sociologique) de communautés que de conceptions philosophiques à part entière. Pour notre argumentation, il faut cependant noter que le modèle médical et le modèle psychologique supposent l'existence de variables latentes. Le premier énonce clairement un substrat « physique », étiopathogénique sous-jacent aux phénotypes ; le second suppose un ensemble de construits psychologiques (par exemple, des « processus » dans l'approche processuelle), adoptant souvent une approche plus dimensionnelle, comme c'est le cas pour la HiTOP. Autrement dit, la variable latente représentée par l'éventuelle « cause physiopathologique » du DSM (renvoyant au modèle médical) ne fait qu'être substituée par la variable latente représentée par la dimension statistique (renvoyant au modèle psychologique).

Enfin, d'un point de vue méthodologique, le déploiement de la HiTOP nécessite la réalisation d'études longitudinales, qui soutiendront également la validation de ses construits, à l'image des essais sur le terrain (les *field trials*) réalisés dans le cadre des différentes versions du DSM (Clarke et al., 2013). Ils pourraient également avoir à s'harmoniser avec les RDoC. Au-delà de l'évolution des construits, dans une visée longitudinale, il semblerait important que chaque construit de la HiTOP prenne en compte l'aspect évolutif des pathologies, se rapprochant alors des modèles de stadification (ou *staging*), qui proposent des construits dimensionnels évoluant en fonction du pronostic d'un trouble donné (McGorry & Hickie, 2019). À ce propos, des propositions de révisions de la classification ont déjà eu lieu, avec la volonté de standardiser le processus de révision (Forbes, 2020).

3. Enjeux inhérents à la dimensionnalité

Une classification comme le DSM identifie un seuil à partir duquel une condition devient un trouble. Au-dessus de ce seuil, un ensemble de symptômes devient suffisamment préjudiciable pour constituer une préoccupation clinique (nécessitant potentiellement un traitement). Du fait même de son caractère dimensionnel, la HiTOP ne peut fournir un tel seuil. Certes, elle permet de s'éloigner d'un point de coupure entre le normal et la pathologique, qui serait unique pour tous les patients. Elle différencie des degrés d'expression d'une dimension, plutôt que des catégories aux limites nettes (Kotov et al., 2017; Zimmerman et al., 2015). La HiTOP différencie des profils de patients évoluant selon la sévérité symptomatique qu'ils présentent. En ce sens, les états sous le seuil diagnostique du DSM (par exemple, les ultra-haut risques de psychose) peuvent être mieux compris dans le cadre de la HiTOP que dans le cadre du DSM. En termes d'intensité de symptôme et donc de sévérité diagnostique et de prise en charge, les frontières sont « floues » entre patients. Contrairement au DSM où tous les patients d'une même catégorie possèdent un seuil de détresse équivalent, avec potentiellement des symptômes de nature différente, les patients d'une

même dimension de la HiTOP possèdent des intensités différentes d'expression des symptômes. Cet avantage apparent pourrait cependant introduire un risque de médicaliser des conditions non pathologiques, risque sur lequel il s'agit de rester particulièrement vigilant. Partant de ce constat, dans la HiTOP, ce qui est gagné concernant la précision descriptive de l'intensité symptomatique pose de nouveaux problèmes autour de la difficile question de la pathologisation et de la médicalisation.

4. Enjeux d'utilité clinique

L'utilisation des dimensions en pratique clinique a été montrée dans d'autres spécialités médicales. Fulford et Handa donnent l'exemple de la chirurgie vasculaire qui, en tant que spécialité relativement nouvelle (la *Vascular Surgical Society* de Grande-Bretagne et d'Irlande a été fondée en 1966), a pu se développer grâce à une approche résolument catégorielle (Fulford & Handa, 2018). Mais ce caractère n'empêche pas cette discipline d'adopter une conception dimensionnelle : par exemple, le traitement de l'anévrisme aortique a progressé grâce aux études quantitatives sur le risque relatif de décès en fonction du diamètre (mesure dimensionnelle) de l'anévrisme (car le risque de rupture augmente à mesure que l'anévrisme se dilate). La combinaison d'approches catégorielles et dimensionnelles semble donc fructueuse dans le cadre de cette discipline médicale, même si son objet d'intérêt (l'anévrisme) reste une entité définie de manière catégorielle. Comme le souligne Fulford, toutes les sciences progressent grâce à la quantification, « mais le progrès par la quantification s'est généralement fait par addition [aux modèles catégoriels] et non par substitution » (Fulford, 2008). Si la HiTOP constitue une avancée quantitative pertinente pour la psychopathologie, il semblerait que ce soit comme un ajout par rapport aux modèles existants plutôt que comme un déplacement du cadre nosologique vers un nouveau système. Cependant, nous soutenons que le fait de rendre dimensionnel des construits distincts les uns des autres au sein d'une classification (comme s'il s'agissait de catégories distinctes les unes des autres) rend complexe l'utilisation de la HiTOP. Par

exemple, les spectres d'extériorisation et d'intériorisation ne saisissent pas les patients présentant un état mixte, une forme de bipolarité mêlant des symptômes d'épisodes maniaques et dépressifs. De plus, la distinction horizontale entre les dimensions empêche de penser certains tableaux cliniques imbriquant plusieurs facettes. Par exemple, notamment en raison de cette difficulté, la HiTOP n'intègre pas les troubles neurodéveloppementaux dans sa structure actuelle. Comme l'affirmait Ghaemi, il est à craindre que cette revendication de dimensionnalité ne se limite finalement qu'aux troubles de la personnalité (Ghaemi, 2018). Afin de mieux intégrer la HiTOP à la pratique clinique, Tyrer proposait d'étudier en parallèle deux classifications, afin de permettre leur comparaison. Un système « dominant », catégoriel ou dimensionnel, ne serait à considérer que secondairement à cette comparaison. Amener subtilement la HiTOP à être utilisée en pratique clinique nécessite de considérer le relatif conservatisme des cliniciens pour la classification qu'ils ont l'habitude d'utiliser (le DSM)⁸³. En plus de ce conservatisme qui empêche la HiTOP de s'intégrer pleinement à la pratique, la classification bute sur certaines questions fondamentales que pose la clinique : à quel moment un comportement normal doit-il être médicalement considéré ? À partir de quand doit-on traiter une condition (comprise de manière dimensionnelle) ? À quel moment est-il raisonnable de considérer des troubles cooccurrents ou des comorbidités, comme c'est souvent le cas avec l'anxiété et la dépression ?

Au travers de ce genre de questions, il est possible de s'apercevoir combien il est difficile d'échapper aux approches catégorielles. De nombreuses décisions de prises en charge sont catégorielles, comme par exemple une décision d'hospitalisation. Toutefois, les médecins se préoccupent rarement de « définir ce

⁸³ L'association d'une perspective dimensionnelle et d'une perspective catégorielle permet de concevoir un cadre de travail clinique dans lequel la catégorisation ne serait qu'un des deux axes, avec l'approche plus individualisante et raffinée offerte par la dimensionnalité. Ces deux axes seraient complémentaires et pas spécialement chronologiques même si, en recherche, ce cadre de travail semblerait pertinent, en isolant d'abord les dimensions puis en recherchant les causes de manière catégorielle, dans un second temps. D'ailleurs, en termes d'articulation entre le dimensionnel et le catégoriel, rappelons que la CIM-11 a longuement travaillé à intégrer des constructions dimensionnelles dans la classification des catégories psychiatriques. Par exemple, on y retrouve l'incorporation d'une classification dimensionnelle des troubles de la personnalité. De même, les sous-types de schizophrénie de la CIM-10 (par exemple, paranoïdes, hébéphréniques, catatoniques) ont été remplacés par un ensemble de classements des symptômes (par exemple, symptômes positifs, symptômes négatifs, symptômes cognitifs).

qu'est un trouble médical et s'attardent plutôt à diagnostiquer et à traiter du mieux qu'ils le peuvent des patients singuliers », selon les termes de Kendell (Kendell, 1975), repris par Spitzer et al. dans leur article sur le *Research Diagnostic Criteria* en 1978 (Spitzer et al., 1978). En effet, les cliniciens ne conduisent pas nécessairement leur prise en charge en fonction du diagnostic catégoriel (Baldwin & Kosky, 2007; Mohamed et al., 2008; Taylor et al., 2016). Ils se concentrent plutôt sur les symptômes, la souffrance et les plaintes (nous reviendrons sur ce point dans la seconde partie de la thèse). Comme le montre un certain nombre d'études récentes portant sur le raisonnement clinique (par exemple (Hopwood et al., 2020)), la structure de la HiTOP pourrait être plus intuitive pour le clinicien. Enfin, l'absence de catégories a également des conséquences sur les plans sociaux et administratifs. Les gouvernements des États membres de l'OMS ont intégré la CIM et son approche catégorielle dans les processus cliniques et politiques, que ce soit en rapport avec la couverture et le remboursement des soins de santé, les services sociaux ou les prestations d'invalidité (World Health Organization, 2017). De plus, une telle classification catégorielle est utile pour assurer la continuité des données de santé et l'application des politiques de santé publique, dans la coordination avec la classification des autres troubles et dans la recherche classique portant sur les mécanismes connexes étiologiques et la comorbidité. Le groupe de travail œuvrant à la « traduction clinique » de la HiTOP a développé un tableau de concordance HiTOP – CIM-11 afin d'aider à relier ses dimensions aux codes de la CIM (Reed, 2018). Cependant, dans sa structure, la HiTOP est loin de ressembler aux catégories de la CIM-11 et leur adéquation nécessite un réel travail de traduction – qui ne semble pas parfaitement réalisable (en raison notamment de la hiérarchisation). Pour lui offrir une entière intégration dans le paysage classificatoire, en plus de répondre aux divers enjeux cliniques, il faudra que la HiTOP réponde également à des enjeux sociaux, administratifs et politiques.

Conclusion

En conclusion, la classification dimensionnelle et hiérarchique de la psychopathologie comporte un certain nombre d'atouts et d'enjeux. L'apport potentiel et original d'une telle classification est de limiter les frontières arbitraires, la comorbidité artéfactuelle, les symptômes de faibles spécificité et l'hétérogénéité diagnostique des approches catégorielles. En répondant à certaines de ces contraintes, la HiTOP réussit pour une part à se positionner aux frontières d'une classification historique comme le DSM et d'un projet neurobiologique comme le projet RDoC. Cependant, les défis de la HiTOP tiennent à sa validité, au renforcement de sa méthodologie et à son utilité pour le raisonnement clinique, sa capacité à fournir des seuils diagnostiques et thérapeutiques et son intérêt pour la compréhension des mécanismes sous-jacent aux troubles psychiatriques. Ces défis nous permettent de mettre en avant que la HiTOP, structurée autour de variables latentes et à l'instar des classifications décrites dans les chapitres précédents, tombe dans un écueil relativement similaire à ces dernières. Comme nous le verrons dans la seconde partie de cette thèse, cet ancrage de la définition d'une espèce psychiatrique à une variable latente semble être au cœur des difficultés à réduire l'écart entre les fonctions cliniques et non-cliniques.

TRANSITION

Parce qu'elles s'intègrent à une approche par la catégorie ou la dimension, les classifications présentées dans la première partie de cette thèse possèdent une caractéristique commune : la notion de variable latente⁸⁴. Les classifications psychiatriques sont structurées sur la base de catégories ou de dimensions qui constituent des variables latentes. La difficulté à identifier de telles variables latentes pour les troubles psychiatriques pourrait être à l'origine d'un grand nombre de désagréments retrouvés dans la clinique psychiatrique, en recherche et, par extension, en philosophie de la psychiatrie. C'est justement cette notion de variable latente qui nous servira de point de bascule entre l'analyse des conceptions et classifications fondées sur une approche par la catégorie ou la dimension et l'analyse d'une conception fondée sur une approche par le symptôme. **L'approche par le symptôme récuse en effet toute notion de variable latente pour définir une espèce psychiatrique.**

Nous cherchons dans ce travail à démontrer une thèse relativement simple : s'il existe différentes classifications et donc plusieurs conceptions en psychiatrie, c'est parce que le désir de répondre à des fonctions à la fois cliniques et non-cliniques a constitué un des ressorts essentiels du développement de la discipline, depuis au moins la parution du DSM-III. Il existe ainsi une forte tension en psychiatrie liée à la volonté de répondre à une variété d'objectifs, tension qui aurait conduit à la multiplicité des classifications. Nous tentons de répondre à cette tension en proposant de nous décaler de l'approche par la catégorie ou la dimension, telle qu'elle a été développée au sein du DSM et au sein d'autres projets ayant tenté de sortir de ses impasses, comme le projet RDoC ou la HiTOP – toutes trois témoins de conceptions historiquement ancrées, neuroscientifiques et statistiques de la psychiatrie. L'approche par le symptôme sur laquelle que nous développons refuse de s'appuyer sur la notion de variable latente pour définir une espèce psychiatrique. Elle semblerait plus fructueuse et pertinente pour la pratique clinique, comme nous cherchons à le montrer dans la seconde partie de cette thèse. Nous verrons en particulier

⁸⁴ Rappelons une fois encore que la variable latente se définit comme une variable psychiatrique qui ne peut être déduite qu'indirectement à partir d'autres variables observables ou mesurables. La variable latente peut être ou non une cause commune. Une variable latente sert à regrouper plusieurs symptômes en une seule entité sous-jacente.

comme un tel réseau pourrait éclairer les raisons de la multiplicité de classifications dans le paysage psychiatrique contemporain, qui résulterait d'une forme d'« impatience » conceptuelle et de « sauts » d'une classification à l'autre alors qu'il faudrait plutôt en étudier les changements locaux un à un, par glissements sémiologiques.

L'appropriation d'une approche par le symptôme pour la pratique clinique ne nécessite pas d'abandonner toute approche par la catégorie / dimension. Comme nous l'avons largement détaillé dans la première partie de cette thèse, la notion de variable latente est particulièrement utile pour répondre à des exigences politiques ou administratives, communiquer avec d'autres professionnels de santé ou soutenir l'enseignement de la psychiatrie⁸⁵. Par exemple, les catégories et dimensions – et la notion de variable latente sur laquelle elles reposent – peuvent avoir des implications en termes de choix de prise en charge, de prédiction d'évolution ou d'évaluation pronostique sur le plan épidémiologique et populationnel. Cependant, ces implications n'ont avant tout de pertinence qu'en dehors de la pratique clinique, définie de manière restreinte comme l'entretien clinique psychiatrique en face-à-face. Une approche par la catégorie / dimension semble nécessaire en complément d'une approche par le symptôme. C'est justement une telle approche par la catégorie qu'offre la définition issue de l'analyse de la dysfonction préjudiciable (*Harmful Dysfunction Analysis* – HDA), qui conçoit un trouble psychiatrique comme la dysfonction, préjudiciable, de certains mécanismes internes pour exécuter une fonction physiologique naturellement sélectionnée. Dans cet exemple de définition générique, la dysfonction correspond à l'échec d'une caractéristique à exécuter une fonction physiologique naturelle pour laquelle elle est biologiquement conçue par des processus évolutifs ; le préjudice est défini comme une détresse, vécue par une personne, ou une incapacité, décrite par une société et conformément aux valeurs sociales d'un temps et d'un lieu donné (Wakefield, 1992). Nous décrirons plus largement l'HDA dans la seconde partie de la thèse.

⁸⁵ Bien entendu, ce n'est pas tant la « notion » de variable latente qui est directement utile, mais bien sont application permettant de constituer des catégories et des dimensions.

Le problème que nous avons posé dans la première partie de cette thèse est que les diverses tentatives d'opérationnalisation des catégories ou dimensions ont pris comme point de départ le fait que les classifications devraient (entre autres) servir à la clinique. Nous y avons souligné cet écart entre promesse d'utilisation clinique et utilisation clinique effective. Ces promesses ne sont d'ailleurs pas toujours explicites au sein des classifications que nous avons présentées : le DSM avait été proposé avant tout comme un manuel pour la recherche, mais il doit aujourd'hui sa pérennité à ses promesses d'utilisation clinique ; le projet RDoC a certes été présenté comme un programme neuroscientifique, mais il doit sa crédibilité et sa légitimité à son nécessaire impact pour la clinique ; la HiTOP, elle, a clairement été proposée pour remplir des objectifs cliniques. Cependant, pour n'en donner qu'un exemple, les résultats d'enquêtes sur le terrain montrent bien qu'un grand nombre d'utilisateurs du DSM jugeait que ses catégories avaient une faible utilité pour « choisir un traitement » et « évaluer le probable pronostic » (First et al., 2018). Ainsi, si les classifications doivent être utiles à la clinique et qu'elles reposent sur une notion (de variable latente) qui ne « parle » pas aux cliniciens, ne devrait-on pas penser une conception de la psychiatrie plus intuitive pour eux ? D'autres auteurs ont commencé à penser ce genre de décalage, comme l'écrivait Mario Maj :

« Il devient de plus en plus évident que l'utilité des catégories diagnostiques en psychiatrie a été surestimée. Ces catégories ont été initialement chargées d'implications en termes de traitement et ont été prospectivement chargées de renvoyer à une étiologie et / ou pathogenèse spécifique, en complète analogie avec les autres branches de la médecine » (Maj, 2018).

En nous décalant des définitions qui se fondent sur les variables latentes, nous serons amenés à proposer une définition d'un trouble psychiatrique plus fructueuse pour la clinique⁸⁶. Évaluer la sévérité d'un

⁸⁶ D'une certaine manière, en diminuant la primauté de la notion de variable latente, ce travail s'inscrit dans la lignée de l'importante littérature critique du modèle médical en psychiatrie. Cependant, il n'emploie pas les mêmes arguments, en s'appuyant entre autres sur les conceptions et classifications psychiatriques, la définition de ce qu'est une espèce psychiatrique, la notion de fonctions (non-)cliniques – tout en proposant une approche complémentaire et non exclusive à la définition d'une espèce psychiatrique cohérente avec ce modèle médical.

tableau clinique, recueillir les antécédents personnels et familiaux, prendre en compte des facteurs périnataux et environnementaux, de développement et de neurodéveloppement, de risque psychosocial, de facteurs biologiques et génétiques mesurables, de traits de personnalité et de scores de psychométrie, de fonctionnement neurocognitif et de cognition sociale, de ressources adaptatives et de résilience, de signes neurologiques légers, d'addiction et de dépendance, d'expositions environnementales et bien sûr de symptômes propres à chaque patient, semble être de véritables gageures pour une unique définition générique reposant sur la notion de variable latente – comme l'HDA. Face à cet ensemble de facteurs nécessaires à prendre en compte dans la pratique clinique, les cliniciens ont besoin d'une définition claire et explicite et en même temps intuitive sur laquelle s'appuyer lorsqu'ils effectuent leur entretien clinique. Nous chercherons à expliciter la définition (implicite) sur laquelle repose l'intuition clinique⁸⁷. Dans la suite de cette thèse, nous soutiendrons que l'accent mis sur les catégories et dimensions dans les recommandations cliniques, les interventions pharmacologiques, le codage administratif ou la pédagogie a minimisé pour une part l'importance d'une définition utile pour le clinicien, dans sa clinique quotidienne et pour un patient donné. Au travers de multiples exemples cliniques, dans la partie finale de cette thèse, nous présenterons une telle définition, dite « théorico-clinique ». Cette définition, renvoyant bien à une espèce psychiatrique, ne sera pas fondée sur la notion de variable latente.

Sur le plan méthodologique, afin de proposer une telle définition utile pour la clinique, nous nous pencherons sur la question du symptôme. Nous nous appuyerons sur les outils fournis par la théorie des réseaux de symptômes pour avancer vers une formulation d'une définition théorico-clinique pour le clinicien. Pour n'en donner qu'un exemple, nous montrerons qu'une telle définition renforce la capacité des cliniciens à rechercher rapidement d'autres symptômes à partir d'un premier symptôme recueilli (favorisant les « heuristiques » cliniques), par exemple en s'appuyant sur des symptômes-ponts.

⁸⁷ Ce projet est en quelque sorte à l'image du projet de définition d'une maladie de Christopher Boorse (1975), qui cherchait à éclairer un implicite dans les pratiques et théorisations médicales.

Une définition potentiellement utile au clinicien devrait-elle être insérée au sein d'une classification ? Nous verrons que tout l'enjeu d'une telle définition sera d'éviter de créer ou proposer une nouvelle classification psychiatrique. La définition proposée sera à la fois stable et manipulable par le clinicien. L'analyse des fonctions cliniques et non-cliniques des classifications présentées précédemment nous sera utile pour étudier la complémentarité d'une définition fondée sur le symptôme et d'une définition fondée sur une variable latente. Nous verrons qu'au moins deux approches (et donc deux définitions) devraient cohabiter en fonction de leur contexte d'utilisation⁸⁸. Le **Tableau 2** permet d'illustrer les caractéristiques d'une telle approche par le symptôme relativement à l'approche par la catégorie.

Tableau 2 – Équivalence des termes renvoyant à l'approche par la catégorie / dimension et à l'approche par la symptôme. Nous nous appuyons sur l'exemple de l'analyse de la dysfonction préjudiciable. DSM : *Diagnostic and Statistical Manual of Mental Disorders* ; HiTOP : *Hierarchical Taxonomy Of Psychopathology* ; RDoC : *Research Domain Criteria*. Nous présenterons par la suite ce que nous nommons déjà ici une « espèce complexe ».

Niveau de description	Approche par la catégorie / dimension	Approche par le symptôme
Entité de base	Variable latente	Réseau de symptômes
Espèce psychiatrique	Espèce diagnostique	Espèce complexe
Conceptions	DSM, HiTOP ou projet RDoC (« modèle médical »)	Prioritairement axée sur la pratique clinique (plus proche du « modèle psychologique »)
Exemple	Le trouble de la communication sociale pragmatique est défini selon une dysfonction dans la communication sociale entraînant un préjudice pour la	Les symptômes de « suivi des règles de la conversation », « d'adaptation de la communication au contexte » ou de « mauvaise compréhension de

⁸⁸ À noter que, volontairement, nous n'avons pas cherché à fournir une définition spécifique de ce qu'est une espèce psychiatrique au sein de chacune des classifications précédentes, le DSM, le projet RDoC et la HiTOP. Nous resterons sur l'idée que toutes se fondent sur la notion de variable latente.

	qualité des interactions de l'individu (exemple de l'analyse de la dysfonction préjudiciable).	l'implicite » constituent des symptômes préjudiciables s'influçant mutuellement.
--	--	--

PARTIE II

DE L'APPROCHE PAR LA CATÉGORIE À L'APPROCHE PAR LE SYMPTÔME

La psychiatrie doit répondre à des fonctions cliniques et non-cliniques. Diverses conceptions et classifications ont été proposées pour tenter de faire face à ces différentes fonctions. Chacune de ces conceptions est sous-tendue par au moins une définition de ce qu'est une espèce psychiatrique. Comme nous l'avons vu dans les trois premiers chapitres, ces définitions sont fondées sur la notion de variable latente. Cet ancrage à la notion de variable latente semble comporter un certain nombre de limites, notamment pour répondre aux fonctions cliniques de la psychiatrie.

Dans ce chapitre, nous soutenons que les fonctions cliniques doivent s'appuyer sur une définition de la psychiatrie qui ne se fonde pas sur la notion de variable latente. Les fonctions cliniques nécessitent l'adoption d'une approche par le symptôme. La conception des réseaux de symptômes fournit un certain nombre d'outils permettant de discuter cette approche. Nous prenons pour exemple l'HDA en tant qu'approche par la catégorie pour offrir un contraste à notre argumentation.

Afin d'étudier l'intérêt d'une approche par le symptôme pour la psychiatrie, nous interrogeons en premier lieu la place du symptôme psychiatrique. Pour cela, nous identifions et déconstruisons deux présupposés, promus en philosophie de la psychiatrie, qui auraient empêché de penser une telle approche par le symptôme. Le premier présupposé concerne la conviction que les cliniciens gagneraient à concevoir les troubles psychiatriques comme des entités stables aux limites claires. Nous soutenons en retour une définition d'une espèce psychiatrique dans laquelle les symptômes rendent compte de mécanismes activés, en interactions mutuelles, au sein d'un noyau central prototypique. Le deuxième présupposé concerne la manière dont la philosophie de la psychiatrie a abordé la question du préjudice. Nous proposons que les cliniciens perçoivent principalement les préjudices à travers des symptômes, en nous appuyant sur le fait que le préjudice est mieux décrit au niveau du symptôme. Nous montrons de quelle manière de tels symptômes préjudiciables impliquent l'altération de mécanismes (biologiques).

L'identification et la déconstruction de ces deux présupposés nous conduisent à proposer une définition théorico-clinique d'une espèce psychiatrique utile au clinicien (nommée « espèce complexe »). Nous proposons la définition suivante d'un trouble psychiatrique pertinente et intuitive pour la pratique clinique : « un ensemble de symptômes appartenant au noyau central prototypique de la psychiatrie, organisés en réseau, impliquant au moins un préjudice et dont l'ensemble des préjudices dépasse le seuil de signification socialement accepté selon les valeurs sociales admises dans une société pour une condition donnée, chacun de ces symptômes étant lié à un mécanisme activé. Un tel réseau de symptômes est figé dans un état qui ne permet pas la diminution spontanée du préjudice relié à ces symptômes sans intervention thérapeutique ».

En parallèle de cette proposition, nous adoptons un principe de conservatisme prudentiel qui n'exclut pas l'existence d'autres définitions du trouble psychiatrique. La définition fondée sur l'HDA, sous-tendue par la notion de variable latente, pourrait en particulier servir de hub épistémique pour l'ensemble des fonctions non-cliniques que requiert la psychiatrie.

Introduction

Il existe une multiplicité de conceptions en psychiatrie. Ce travail de thèse ne cherche pas à en fournir une liste exhaustive ou une synthèse. Il ne tente pas non plus d'en apporter une analyse philosophique – par exemple, en tant qu'ils seraient des « programmes de recherche » ou des « matrices disciplinaires » – ni une analyse au prisme de la clinique – en opposant ou en distinguant les représentations de la psychiatrie par les psychanalystes, les neurobiologistes ou les phénoménologues. Dans les chapitres précédents, nous avons plutôt étudié les conceptions psychiatriques au travers de certaines de leurs instanciations concrètes que sont les classifications. Nous avons émis l'hypothèse que cette multiplicité de conceptions et classifications psychiatriques serait liée à leur volonté de répondre à la fois à des fonctions cliniques (diagnostiques, pronostiques et thérapeutiques) et non-cliniques (par exemple, institutionnelles, politiques ou financières). Or, cette volonté de résoudre la tension entre fonctions cliniques et non-cliniques a conduit les conceptions et classifications psychiatriques à un grand nombre de difficultés, abordées dans les chapitres précédents. Le fait que chacune de ces conceptions et classifications adopte une approche par la catégorie / dimension aurait pu conduire à ces difficultés, comme nous l'analyserons dans ce chapitre. Nous répondons à cette tension entre fonctions cliniques et non-cliniques en assumant cet écart, en proposant de séparer deux définitions d'une espèce psychiatrique : l'une pour répondre aux fonctions cliniques et l'autre pour répondre aux fonctions non-cliniques.

La première est une définition générique. Elle sert à déterminer quelles conditions doivent être considérées comme pathologiques, afin de répondre à des objectifs administratifs, financiers, politiques, de recherche, d'organisation sociale du système de santé, de politiques de soins et de recommandations de bonnes pratiques (Frances 2013 ; Spitzer 1983). Nous prenons l'exemple de la HDA, qui fournit une définition ayant un fort potentiel épistémique et pragmatique pour servir de *hub* épistémique (Keuck,

2011 ; Wakefield 1992). Nous définissons un *hub* épistémique comme tout outil scientifique (ici, une définition) qui centralise des perspectives de connaissances issues de multiples sources. En servant de « concentrateur » pour différents types d'informations, ces hubs peuvent faciliter les échanges, la communication et les controverses entre différentes institutions (par exemple, les institutions médicales et les communautés de patients) ou différentes disciplines (par exemple, la médecine et les neurosciences théoriques).

La seconde définition est dite « théorico-clinique ». Elle serait utile à la pratique clinique psychiatrique (Krueger, 2015). Elle considère les troubles psychiatriques au travers du regard du clinicien, qui raisonne et prend des décisions au niveau clinique, pronostique et thérapeutique. Pour ces décisions, qui concernent entre autres la réorientation, le conseil, l'accompagnement, la prévention des risques, l'élimination des diagnostics différentiels, le clinicien a besoin d'une définition manipulable, précise et intuitive (nous reviendrons sur ces termes). Par la suite, nous donnons une définition originale de ce qu'est une telle espèce psychiatrique, nommée « espèce complexe ». Nous précisons déjà ici que cet intérêt destiné à fournir au clinicien une définition spécifique à sa pratique ne prétend en rien admettre l'existence d'une « exclusivité » de la clinique par rapport aux autres champs cités précédemment (Demazeux, 2019).

Notre point de départ consiste à adopter une approche fondée sur le symptôme, utile pour le clinicien. Nous nous décalons du modèle médical pour aborder un modèle plus psychologique⁸⁹. Nous contrasterons cette approche avec l'approche par la catégorie / dimension exemplifiée par la HDA. Nous proposons de discuter deux présupposés trouvés en philosophie de la psychiatrie, qui constituent les deux sections principales de ce chapitre.

⁸⁹ Cependant, ce terme de « modèle psychologique » ne correspondant pas à notre approche, nous ne l'utilisons pas ici. En effet, il se rapporte tantôt à une part du modèle biopsychosocial de Engel, tantôt à des approches cognitives (processuelles) et est régulièrement utilisé de manière triviale pour faire référence à tout ce qui touche à la psychologie en tant que discipline. Nous avons choisi ici de parler de modèle « plus » psychologique pour bien mettre en avant l'existence du continuum entre ces modèles (médical et psychologique) et signifier que nous nous décalons d'un modèle fondé sur l'identification d'une variable latente.

Le premier présupposé concerne la croyance (souvent implicite) que les cliniciens perçoivent les troubles psychiatriques comme des entités aux frontières claires, sous-tendues par une dysfonction unique. En d'autres termes, il s'agit de dénoncer une forme de réalisme naïf fondé sur la croyance en un naturalisme des espèces sous-jacentes à une variable latente, tel qu'il peut être pensé par les cliniciens qui utilisent traditionnellement une approche par la catégorie / dimension. C'est la déconstruction de ce premier présupposé qui ouvre la voie à la réflexion en termes d'approche par le symptôme. Le deuxième présupposé concerne la manière dont la philosophie de la psychiatrie a abordé la question du préjudice. Nous supposons que les cliniciens perçoivent les préjudices principalement à travers le recueil et la lecture des symptômes. L'analyse de ces deux présupposés nous amène à reconsidérer la notion de défaillance d'un mécanisme (sur la base du présupposé 1) et de préjudice (sur la base du présupposé 2) – en miroir des deux parties de la HDA – et nous conduit à proposer une définition de ce qu'est une espèce psychiatrique utile, intuitive, pour le clinicien.

Présupposé I : Variables latentes et mécanismes locaux

Nous avons vu dans les chapitres précédents que les cliniciens et chercheurs, quelle que soit l'adoption de telle ou telle classification, tendent à employer une conception des troubles psychiatriques basée sur des variables latentes. En considérant l'exemple de la HDA, la dysfonction (part factuelle de cette analyse) constitue une telle variable latente. Il s'agit en effet d'une entité (fonctionnelle) non observable expliquant l'agrégation de symptômes au sein d'un trouble. Par exemple, le trouble bipolaire serait en partie considéré comme un trouble en tant qu'il est le produit d'une dysfonction (préjudiciable).

Dans cette section, nous montrons que la conception du clinicien de ce qu'est un trouble psychiatrique est mieux renseignée par un réseau de symptômes activant des mécanismes locaux plutôt que par l'analyse d'une dysfonction sous-jacente à une espèce psychiatrique⁹⁰.

1. Approche par le symptôme

a. Influence causale mutuelle

L'expérience clinique, étayée par des recherches empiriques, a montré que les cliniciens ont tendance à penser, intuitivement, en termes d'influences causales mutuelles entre symptômes⁹¹ (Kim et Ahn, 2002). Par exemple, le clinicien recueille un symptôme de tristesse en observant une réaction physique comportementale (les pleurs) faisant suite à une cognition négative (le souvenir d'un événement) influencée elle-même par une métacognition (la culpabilité de repenser à ces souvenirs) en lien avec l'existence d'expériences vécues (le sentiment d'abandon). Ces interactions entre symptômes peuvent être circulaires (par exemple, les pleurs augmentent eux-mêmes la perception d'une cognition). Dans cette succession d'influences mutuelles, identifier une « dysfonction », au sens de la HDA, semble particulièrement problématique. De plus, une telle intuition est au fondement des thérapies cognitivo-comportementales : les cinq colonnes de Beck, cadre de travail qui sous-tend ces thérapies, reposent en

⁹⁰ Notons que, dans le débat sur les espèces psychiatriques, il semble nécessaire de bien différencier la question du normal et du pathologique (par exemple : « Les symptômes appartenant à la catégorie de la dépression constituent-ils, ensemble, un trouble psychiatrique ? ») de la question de la dimensionnalité (par exemple : « Quel degré de tristesse est présent chez ce patient présentant un tableau clinique renvoyant à la catégorie de dépression ? ») et de celle du diagnostic différentiel (par exemple : « En quoi ces symptômes appartenant à la catégorie de la dépression se distinguent-ils de ceux de l'anxiété généralisée ? »). Dans ce chapitre, à partir de la notion de « symptômes préjudiciables », que nous présenterons, nous cherchons à répondre à la première de ces questions, à savoir la différenciation entre normal et pathologique.

⁹¹ Comme nous l'avons déjà souligné en introduction, par commodité, nous regroupons les symptômes, les signes, les facteurs de risque, les pensées, émotions, comportements, expériences somatiques et éléments sociaux et contextuels sous le terme de « symptômes ».

effet sur l'identification d'une situation, qui elle-même influence une émotion via des pensées automatiques. Cette émotion conduit alors à des réponses comportementales et à des conséquences qui renforcent les pensées automatiques lorsque des situations similaires se présentent (Beck 1993). Dans ces modèles et thérapies associées, les symptômes s'influencent mutuellement.

Un tel réseau de symptômes en interaction mutuelle est constitué de nœuds (les symptômes, ou autres variables environnementales, ou marqueurs biologiques) et de connexions entre ces nœuds (Borsboom 2008). Les symptômes possèdent un rôle causal particulier dans l'approche réseau, puisque chacun contient une part de la causalité de l'ensemble du réseau. De fait, les relations entre symptômes et la structure générale du réseau comportent une importance particulière dans cette approche par le symptôme. Sur le plan clinique, les troubles psychiatriques peuvent alors être considérés comme des « cercles vicieux de problèmes causaux » (Stein 2021). Sur le plan ontologique, les troubles psychiatriques peuvent être appréhendés comme des clusters de propriétés homéostatiques (Boyd 1991, 1999 ; Casali 2021 ; Kendler et al. 2011 ; Kendler 2008 ; Wakefield 2012 ; Westen 2012) dont les composants sont des éléments cliniques, phénotypiques (**Annexe 2 – « Approche réseau et clusters de propriétés homéostatiques »**) – par opposition aux composants biologiques, endotypiques (même si la présence de marqueurs biologiques n'est pas exclue de tels réseaux).

b. Abandon de la variable latente

Nous avons défini une variable latente comme un élément non observable (ici, dans la pratique clinique) sous-jacent à un ensemble d'éléments exprimés (ici, les symptômes). Par exemple, dans la pratique psychiatrique contemporaine, la variable latente correspond à la catégorie ou dimension « dépression ». Dans le DSM, la « dépression » ne possède qu'une valeur littérale ou stipulative (ou ostensive – (Sadegh-Zadeh, 2008), c'est-à-dire qu'elle correspond à un terme n'existant pas indépendamment de la définition

à laquelle renvoie ce terme, qu'elle est une dénomination sur laquelle il n'est pas possible de se prononcer sur la nature ou la fonction (**Annexe 2 – « Approche réseau et clusters de propriétés homéostatiques »**). Dans la HiTOP, la variable latente est un construit statistique, puisqu'une dimension extraite d'une analyse factorielle agrège les symptômes (Kotov et al., 2017). Dans la matrice des RDoC, en tant que fonction, la variable latente est un construit psychologique – qui comporte une validité (interne) du fait des unités d'analyse biologiques intégrées à la matrice. Enfin, en prenant l'exemple de la HDA utilisé dans ce chapitre, la variable latente correspond à la « dysfonction » (Putnam, 1975 ; Wakefield, 1999).

Cependant, en pratique clinique, nous soutenons que les ensembles de symptômes ne devraient pas être considérés comme des indicateurs cliniques de variables latentes. Certes, la nature des symptômes (discutée plus loin) dépend des multiples orientations théoriques que peuvent adopter les cliniciens (Demazeux, 2019). Cependant, il ne semble ni nécessaire, ni souhaitable de postuler une variable latente rendant compte de leur agrégation au sein d'une catégorie ou d'une dimension. Nous soutenons que c'est l'ensemble des relations entre les symptômes qui peut le mieux refléter leur présence commune au sein d'un même tableau clinique – et non une supposée variable latente. Ainsi, au lieu de considérer qu'une variable latente agrège des symptômes entre eux, une espèce psychiatrique serait plutôt composée d'un ensemble de symptômes en interactions mutuelles⁹², alors appelée espèce complexe. Par exemple, l'anhédonie conduirait à la tristesse de l'humeur, qui amènerait l'anxiété, conduisant à l'insomnie d'éveil, qui à son tour aggraverait l'anhédonie⁹³. Un symptôme d'insomnie serait présent chez des patients parce qu'il existerait un autre symptôme (comme de l'anxiété) plutôt que parce qu'il existerait une cause commune physiologique sous-jacente à ces deux symptômes. En reformulant cette thèse dans une version plus faible, le raisonnement du clinicien s'appuierait sur ces interactions entre symptômes

⁹² Il sera discuté plus loin que cette définition d'une espèce psychiatrique en tant qu'« ensemble de symptômes en interactions mutuelles » ne devrait pas exclure le fait qu'une espèce psychiatrique puisse aussi être définie comme un « ensemble d'interactions mutuelles entre symptômes ».

⁹³ D'ailleurs, tous ces symptômes n'ont pas le même poids inférentiel en termes diagnostiques, pronostiques, prédictifs ou thérapeutiques. Certains de ces symptômes seront plus ou moins utiles au clinicien en fonction de sa question diagnostique ou thérapeutique.

pour appréhender un tableau clinique (plutôt qu'en leur supposant une dysfonction commune). Comme nous le verrons plus loin, de nombreuses difficultés diagnostiques en psychiatrie, telles que l'hétérogénéité des troubles ou des comorbidités, peuvent être partiellement dépassées en cessant de tenir pour acquise l'existence de variables latentes et en considérant plutôt les interactions entre symptômes (Borsboom et al. 2017 ; Fried et al. 2017 ; McNally 2016).

Bien sûr, le raisonnement selon lequel les variables latentes ne sont pas utiles aux cliniciens ne signifie pas que ces variables doivent être complètement exclues de toute considération en psychiatrie (voir pour cela la possibilité de combiner cette approche par le symptôme avec les variables latentes, dans l'**Annexe 4, « Le spectre d'essentialité »**). Par exemple, certaines études en réseaux de symptômes utilisent des variables latentes comme l'indice de masse corporelle, une variable latente sociodémographique qui influence plusieurs symptômes appartenant au trouble dépressif caractérisé décrit dans le DSM. Ces réseaux sont alors appelés « hybrides ». Cependant, nous soutenons qu'une telle variable latente n'entre pas dans la définition de ce qu'est une espèce psychiatrique⁹⁴. Cet intérêt pour une approche se détournant des variables latentes montre notre intérêt pour le raisonnement du clinicien : nous nous penchons en effet sur un trouble psychiatrique, dans l'objectif d'en proposer une définition utile, afin de déterminer les conditions d'attribution d'un trouble, c'est-à-dire spécifier à partir de quand le clinicien est amené à diagnostiquer et à soigner. Il s'agit cependant de noter que nous discutons ici du raisonnement clinique défini comme l'association de symptômes conduisant à une prise de décision pratique et non d'une prétendue intuition « artistique » que posséderait hypothétiquement le clinicien.

Rappelons également que si la nécessité d'avoir une variable latente unissant les symptômes d'un patient n'est pas primordiale en pratique clinique, elle l'est cependant pour des fonctions non-cliniques et notamment en recherche : les comportements impulsifs d'un ensemble d'enfants, s'ils s'accompagnent

⁹⁴ Encore une fois, nous rappelons que nous ne renonçons pas entièrement à l'idée qu'une variable latente à type de mécanisme physiopathologique commun puisse être utile en psychiatrie clinique. Nous insistons plutôt sur le fait que ce type d'explication basé sur les interactions entre symptômes est plus utile au clinicien.

de symptômes de difficultés attentionnelles et d'hyperactivité, doivent renvoyer à la catégorie (variable latente) de « trouble déficit de l'attention avec hyperactivité », qui elle-même pourrait conduire à un traitement par méthylphénidate (agissant sur les voies dopaminergiques). C'est cette connaissance issue de la recherche qui a permis d'introduire historiquement le méthylphénidate pour ce trouble. En effet, ce traitement n'a été développé en recherche – et n'a pu être déployé réglementairement – qu'avec l'appui d'une catégorie diagnostique. Cependant, il existe une nuance. Ce n'est pas la connaissance de ce trouble qui guide le clinicien vers le traitement : c'est bien l'ensemble des symptômes en interaction. Des exemples similaires pourraient être donnés dans un cadre politique : s'il est impossible d'élaborer un Plan Autisme sur la base de simples agglomérats de symptômes, les différentes déclinaisons de ce plan se fondent sur une variable latente (« l'autisme ») qui permet la discussion entre intervenants sociaux, médicaux, associatifs et politiques.

Renoncer à l'idée de variable latente pour mieux comprendre la pratique clinique nécessite de reconsidérer le modèle médical qui conceptualise les symptômes comme les conséquences d'une entité pathologique sous-jacente. Tous les symptômes se situent alors à un niveau équivalent et interchangeable (Lux et Kendler 2010 ; Schmittmann et al. 2013). En nous fondant sur la nécessité de prioriser certains symptômes par rapport à d'autres au sein de la pratique clinique et sur la difficulté à concevoir qu'une unique variable latente puisse rendre compte de l'agrégation d'un ensemble de symptômes (par exemple, l'anhédonie et l'insomnie qui seraient présents *par hasard ensemble* – pour le dire de manière radicale – chez un même patient), nous soutenons qu'une variable latente (conservons en exemple l'idée d'une dysfonction unique) ne peut pas expliquer l'agrégation de l'ensemble des symptômes d'une espèce psychiatrique.

2. Notion de noyau central

a. Noyau central

Si le clinicien ne cherche pas nécessairement à identifier une variable latente et ne s'appuie pas sur une approche par la catégorie / dimension, comment peut-il trouver une unité dans le tableau clinique d'un individu présentant un ensemble de symptômes en interaction⁹⁵ ? Nous proposons de considérer l'existence d'un noyau central de symptômes au sein du réseau (qui renvoie sur le plan ontologique à un cluster central de propriétés homéostatiques). C'est ce noyau central qui constituerait ce qu'est une espèce psychiatrique.

L'une des principales méthodes utilisées en recherche dans le cadre des modèles de réseau de symptômes correspond à la détermination des symptômes les plus centraux au sein d'un réseau (Bringmann et al. 2019 ; Jones et al. 2019). Un tel noyau de symptômes centraux est un ensemble de symptômes stable dans le temps, entre patients et relativement indépendant des intérêts épistémiques des chercheurs (Amundson et Lauder 1994). C'est donc ce noyau central qui permet de définir ce qu'est une espèce psychiatrique utile pour le clinicien (Zachar 2000).

Le réseau de symptômes d'un patient est également constitué de symptômes périphériques, qui peuvent influencer le noyau central, sans nécessairement entrer dans la définition d'une espèce psychiatrique (Sadegh-Zadeh 2008). Par exemple, un facteur sociodémographique ou un biomarqueur d'inflammation peut influencer le tableau clinique, mais ne correspondra pas au noyau central qui définit l'espèce

⁹⁵ Des auteurs comme Guze auraient répondu à cette question en proposant de différencier les critères monothétiques de critères polythétiques. En effet, les défenseurs des diagnostics fondés sur de tels critères ont tenté de *hiérarchiser* les critères d'un trouble. Cette hiérarchisation structure l'unité du trouble : c'est bien parce que des critères (monothétiques) conditionnent le recueil et le « décompte » d'autres items (polythétiques) – et vice-versa – qu'un diagnostic peut être posé. Plus de quarante ans plus tard, la psychiatrie a évolué et nous ne pouvons plus accepter cette différenciation non fondée sur les données probantes, qui n'est que de faible utilité dans la pratique clinique (Guze, 1978).

psychiatrique (**Annexe 3 – « Définition du noyau central »**). Nous allons détailler quelles conditions de possibilité permettent d'affirmer que certains symptômes sont plus centraux que d'autres.

La notion de noyau central peut être retrouvée dans la littérature en psychiatrie et en philosophie de la psychiatrie sous différentes appellations. Premièrement, on retrouve la notion de noyau central dans la science des systèmes complexes (Cucuringu et al. 2016 ; Ma et Mondragón 2015 ; Rombach et al. 2014). Un système est défini par un ensemble d'éléments liés entre eux par des relations (Bertalanffy 2015). L'un des principaux objectifs de la science des systèmes complexes est d'isoler et d'analyser les propriétés des éléments appartenant à ces systèmes, en supposant que la combinaison de certaines des propriétés du système permettra la compréhension du comportement de système dans son ensemble. Par conséquent, les éléments qui peuvent être considérés comme centraux dans le système sont intéressants à identifier car ils forment le noyau permettant la compréhension du réseau.

Deuxièmement, la notion de noyau central peut également être retrouvée dans la littérature portant sur la théorie des réseaux de symptômes en psychopathologie. Cette théorie réfère notamment au conservatisme (stabilisateur) du principe 2. La théorie des réseaux de symptômes est en effet fondée sur un ensemble de principes permettant la modélisation des réseaux. Les quatre principes qui constituent la théorie des réseaux de symptômes renvoient à la caractérisation du trouble comme réseau de symptômes en interaction (principe 1), à la correspondance entre composants du réseau et symptômes codifiés par l'histoire de la psychiatrie (principe 2), à la nécessité de relations causales directes entre les symptômes (principe 3) et à la non-trivialité de la topologie du réseau, étant donné que certains symptômes sont plus étroitement liés que d'autres (principe 4). Selon le principe 2, « les composants du réseau psychopathologique correspondent aux problèmes qui ont été codifiés comme symptômes au siècle dernier et apparaissent comme tels dans les manuels diagnostiques actuels » (Borsboom 2017). Une conclusion commune des études empiriques sur les réseaux de symptômes est que les symptômes identifiés comme centraux au sein d'un réseau sont finalement le plus souvent des symptômes

cliniquement pertinents (par exemple, il s'agit des symptômes monothétiques, tels que la tristesse de l'humeur ou l'anhédonie) (Beard et al. 2016 ; Bringmann 2021 ; Fried 2015 ; Fried et al., 2016 ; van Borkulo et al. 2015). Une plongée dans l'histoire de la psychiatrie permettrait de montrer que les symptômes centraux tels qu'intégrés aux réseaux selon le principe 2 de la théorie sont ceux décrits par les auteurs ou cliniciens psychiatriques célèbres dans l'histoire de la psychiatrie (par exemple, nous pourrions penser à Schneider pour la schizophrénie).

Troisièmement, la notion de noyau central a été attestée culturellement. Certains symptômes de la catégorie de dépression sont observés dans toutes les cultures (Kleinman 1988 ; Lahey et al. 1996). L'OMS a ainsi explicitement indiqué dans plusieurs de ses rapports la nécessité de fonder les évaluations cliniques des patients sur un ensemble de « symptômes de base » auxquels d'autres, y compris des éléments spécifiques à la culture, peuvent être ajoutés (Sartorius et Organisation mondiale de la santé 1983). Une étude interculturelle de l'OMS a identifié un « tronc commun » de symptômes dépressifs chez 75 % des patients diagnostiqués dépressifs dans le monde : tristesse, perte de l'élan vital, anxiété, tension, perte d'énergie, diminution de l'intérêt et de la concentration et sentiment d'inutilité (Sartorius et Organisation mondiale de la santé, 1983).

Quatrièmement, la notion de noyau central est étayée par des textes issus de la philosophie de la psychiatrie. Par exemple, c'est ce que Rachel Cooper a appelé des « répétables », dans lesquelles « certaines entités dans le monde se ressemblent et se comportent de manière similaire » (Cooper 2012) – un terme qui évite de statuer sur la nature métaphysique d'un trouble. Cooper insiste sur la dimension clinique des répétables, ainsi constitués de symptômes qui peuvent être retrouvés lors de la formulation de cas cliniques (Cooper 2012). Cependant, comme nous allons le voir, la notion de noyau central comporte un grand nombre de caractéristiques qui diffèrent largement des caractéristiques de ces répétables (voir aussi **Annexe 3 – « Définition du noyau central »**).

b. Composants du noyau central

Cliniquement, quels sont les composants d'un tel noyau central ? En pratique, il peut s'agir des symptômes monothétiques des catégories du DSM, comme l'anhédonie, la tristesse de l'humeur, l'hallucination acoustico-verbale ou la désorganisation idéo-affective. Mais la nature de ces symptômes dépend de chaque trouble et plus de 500 études utilisant les réseaux de symptômes en 2023 ont permis d'identifier des symptômes parfois monothétiques, parfois non et parfois n'ayant pas été décrits dans le DSM ni dans aucune classification (mais provenant de certains questionnaires de psychométrie – citons par exemple le « sommeil non-rafraîchissant » qui n'est pas listé dans les classifications mais constitue un des symptômes les plus importants pour les patients reçus en médecine du sommeil (Gauld et al., 2021)).

Sur le plan philosophique, alors que l'ensemble du réseau peut être assimilé à un cluster de propriétés homéostatiques, les symptômes constituant le noyau central du réseau pourraient être considérés comme un prototype de ressemblance (Sadegh-Zadeh 2008). Cette vision se justifie notamment par ses limites floues (Lilienfeld et Marino 1995 ; Sadegh-Zadeh 2001), qui varie en fonction des études empiriques et des progrès scientifiques (par exemple, la découverte d'un nouveau marqueur ou d'une donnée socio-démographique importante perturbera la délimitation du trouble, sans pour autant modifier son noyau central). Le prototype (qui est ici le noyau central) induit des actions chez les membres de la communauté psychiatrique pour soutenir, surveiller, réorienter ou améliorer l'état de la personne présentant un noyau de symptômes en interactions mutuelles. L'une des principales caractéristiques d'un tel noyau central serait ainsi d'inciter à l'action (ce point est abordé en dernière partie du chapitre), dans la mesure où les membres de la communauté partagent des valeurs et des attitudes fondamentales sur la notion de préjudice (voir présumé 2).

Un patient peut avoir certains symptômes d'un noyau prototypique, sans pour autant présenter nécessairement tous les autres. En d'autres termes, seuls certains symptômes d'un noyau prototypique

peuvent être présents chez un patient spécifique. Par exemple, il est rare qu'un patient présente des compulsions, du délire, des hallucinations, de l'anhédonie, de l'anxiété *et* un trouble de la communication pragmatique⁹⁶.

Nous précisons en effet que le noyau central permet d'obtenir un ensemble de symptômes identique pour un ensemble de patients (sur un plan parfois dit « nomothétique » dans les modèles de réseaux de symptômes), permettant de garantir une certaine stabilité à la discipline psychiatrique (qui contiendrait ainsi un nombre fini d'espèces psychiatriques). Les espèces psychiatriques se chevauchent donc, étant donné l'absence de définition claire des conditions individuelles amenées par les patients. Cette « condition » amenée par les patients correspond à un ensemble indéterminé de symptômes, présentés sous différentes formes et récits, pris dans des contextes de vie et des environnements tous différents et non circonscrits ni bornables. Une discussion pourrait être ouverte à ce niveau autour de « ce qui constitue le champ de la psychiatrie » qui – en dehors du champ de la prévention et des politiques d'allers-verts – dépend avant tout du fait que des patients se présentent d'eux-mêmes (ou sont amenés par la société) au système médical⁹⁷.

En ce qui concerne un patient en particulier (sur un plan dit « idiographique »), celui-ci présentera certains symptômes d'une espèce psychiatrique. La présence de seulement certains des symptômes du noyau central relatif à une condition permet d'accepter que le patient soit intégré à un prototype.

Ce qui fait le cœur d'une espèce psychiatrique de type espèce complexe n'est pas la nature du symptôme mais, comme nous le verrons, l'intensité du préjudice par rapport à un seuil (Sadegh-Zadeh 2008 ; Westen 2012). Ce constat est particulièrement important car il permet de postuler l'existence d'un noyau

⁹⁶ En citant ici les différents symptômes potentiels d'un réseau, nous raisonnons de manière agnostique plutôt que transdiagnostique, c'est-à-dire que ce réseau peut comporter des symptômes provenant éventuellement de multiples catégories telles qu'elles sont détaillées dans le DSM, mais ce réseau ne rend pas compte de plusieurs catégories qui se chevaucheraient. L'aspect transdiagnostique nécessite en effet d'accepter l'existence des catégories.

⁹⁷ Il serait possible de reformuler cette affirmation en adoptant une position encore plus normative : c'est au patient et à la société eux-mêmes de décider, en premier lieu (et hors situation de prévention et dépistage), de consulter ou non.

central de symptômes centraux, tout en refusant de faire correspondre un patient donné à une catégorie ou dimension diagnostique.

3. Notion de mécanisme

a. Activations des mécanismes biologiques

Une espèce psychiatrique peut donc être considérée comme constituée d'un noyau central stable : celle-ci est associée à une couronne périphérique de symptômes évoluant au gré des progrès scientifiques ou des programmes de recherche. Une question qui se pose naturellement lors de la description d'un ensemble de symptômes interconnectés est de savoir d'où viennent ces connexions et quelle est leur nature (Cramer et al. 2016). En effet, pouvoir affirmer qu'un symptôme interagit avec un autre nécessite de préciser si cette interaction agit directement au niveau des symptômes, par exemple « l'anxiété » activant le « réveil précoce », ou au niveau des mécanismes locaux responsables de chaque symptôme (en supposant que chaque symptôme est sous-tendu par un mécanisme). Activer ou modifier un mécanisme local responsable de l'anxiété semble pouvoir activer ou modifier un mécanisme local responsable des réveils précoces.

La notion de mécanisme nous semble importante pour mieux comprendre l'approche par le symptôme. Il est donc important de faire la distinction entre le symptôme et le mécanisme qui justifie l'existence de cette manifestation (Machamer et al. 2000 ; Wakefield 2006). Un mécanisme biologique est dit « activé » si le symptôme qui résulte de ce mécanisme est préjudiciable pour le patient (voir présumé 2). En d'autres termes, l'activation est le pendant mécanistique du préjudice qui, lui, se situe au niveau de la pratique clinique. Par exemple, l'activation d'un mécanisme biologique peut entraîner un symptôme (préjudiciable) de tristesse de l'humeur.

Pourquoi est-il si important de poser l'existence de mécanismes activés ? La réponse est de première importance : ceux-ci permettent d'éviter que les troubles psychiatriques ne soient considérés que comme des constructions sociales (Szasz 1960), c'est-à-dire des agrégats de « dénominations », qui n'en ferait qu'un réseau nominaliste sans base biologique (**Annexe 2 – « Approche réseau et clusters de propriétés homéostatiques »**). Ils permettent d'assurer au noyau central d'être *fondé* sur une approche naturaliste (ce qui n'empêche pas de considérer que ces espèces psychiatriques soient elles-mêmes *définies* comme des noyaux centraux de symptômes recueillis et définis sur le plan clinique). L'existence d'un mécanisme activé permet d'éviter que tout réseau d'éléments en interactions mutuelles ne soit considéré comme une espèce psychiatrique (par exemple, une difficulté financière agissant sur la perte d'un bien matériel, qui agit sur les comportements d'achats, etc.). Sans ce recours au mécanisme, tout réseau de symptômes pourrait être considéré comme une espèce psychiatrique pour la discipline. La notion d'activation permet également d'éviter un autre écueil : celui de définir ce qu'est une « altération ». En effet, en discutant de « mécanisme activé » plutôt que de « mécanisme altéré », nous ne nous prononçons pas sur une quelconque « normalité » du mécanisme – normalité qui peut en retour nécessiter de discuter de fonction normale. En convoquant l'activation d'un mécanisme, nous n'avons ainsi pas à discuter ce qu'est une dysfonction par rapport à une fonction (naturelle). Dans notre proposition, nous n'avons ainsi pas besoin de définir ce qu'est une altération.

La relation entre un mécanisme biologique et un symptôme peut être particulièrement difficile à établir en pratique du fait de l'hétérogénéité des relations (causales et associations statistiques) entre clinique et (neuro)biologie (Clementz et al. 2016, 2020 ; Ivleva et al. 2017). Par exemple, dans le spectre de la psychose, il est difficile de dériver des éléments neurobiologiques à partir de la phénoménologie clinique (par exemple, relier un comportement à une activation d'une région repérée en imagerie par résonance magnétique fonctionnelle) (Clementz et al. 2016). L'inverse est également vrai, car un même élément biologique peut laisser penser qu'il comporte plusieurs expressions cliniques (par exemple, la sécrétion de la dopamine agit dans les troubles de l'attention comme dans les dimensions positives et négatives de

la schizophrénie). On parle dans ce cas de multifinalité – un terme classiquement employé dans la littérature psychométrique qui renvoie peu ou prou à la notion de multiréalisabilité en philosophie de l'esprit. Cette difficulté à établir des relations entre symptômes et éléments biologiques ne remet pas en cause l'idée que les symptômes sont potentiellement liés à des mécanismes. En effet, comme cela a été discuté avec l'adoption d'un « essentialisme de la boîte noire » pour la HDA, dans le cadre de ce qu'est une dysfonction (Medin et Ortony 1989 ; Wakefield 1999), le mécanisme impliqué peut être seulement *présumé*, c'est-à-dire non identifié empiriquement mais fortement postulé ou supposé (voir **Annexe 4** – « **Le spectre d'essentialité** », pour mieux comprendre la notion d'identification présumée).

En plus d'être conceptuellement importante, la connaissance ou la supposition d'un tel mécanisme lié à un symptôme peut être nécessaire pour les cliniciens. Ceux-ci peuvent s'appuyer sur l'existence potentielle de ces mécanismes pour obtenir un aperçu clinique de l'évolution et de l'histoire naturelle des troubles. Comme nous allons largement le détailler ci-dessous, en identifiant des mécanismes en interaction chez un patient, le clinicien peut proposer un projet thérapeutique plus spécifique (comme la psychoéducation ou les psychothérapies) ou une surveillance adaptée à court et à long terme. Par exemple, l'identification de mécanismes permet de soutenir le patient dans son parcours de déstigmatisation, notamment en naturalisant certains de ses symptômes (par exemple, en déresponsabilisant pour une part des symptômes d'inattention qui étaient pris pour du refus scolaire chez un jeune patient).

b. Mécanismes à l'œuvre au sein des réseaux

Mais de quel type de mécanisme biologique parle-t-on ? L'intérêt pour la notion de mécanisme s'est accru ces dernières années, comme l'illustre par exemple la littérature portant sur une nouvelle philosophie mécanistique. Cet intérêt prête main-forte à notre thèse, car le fait de s'appuyer sur la notion

de mécanisme plutôt que sur la notion de fonction permet d'éviter les difficultés à définir ce qu'est une fonction. Deux conceptions du mécanisme ont été proposées pour soutenir la définition d'un trouble médical.

La première conception postule que le trouble lui-même devrait être conçu comme un mécanisme (dysfonctionnel) : le trouble s'apparente à l'altération de certains de ces mécanismes internes (Darden 2008 ; Moghaddam-Taaheri 2011). Il s'agit donc de confondre volontairement trouble et mécanisme. Dans cette conception, les troubles sont interprétés comme des mécanismes délétère pour l'organisme dans son intégralité, à l'image d'une machine qui surchauffe lorsqu'un de ses composants tombe en panne (Darden 2008).

Contrairement à cette première conception, une seconde position soutient que les mécanismes constituent des ensembles de variables, structurées en tant que systèmes indépendants d'un autre ensemble (par exemple, de symptômes) désigné comme « trouble ». Les mécanismes sont donc sous-jacents à ces troubles et sont dits « autonomes » (Illari et Williamson 2011 ; Nervi 2010). Il s'agit donc de concevoir le mécanisme comme une variable latente, sous-jacente au trouble (par exemple, constitué de symptômes). Dans cette conception, le trouble n'est pas le mécanisme ; il en est le reflet.

Dans cette thèse, nous acceptons cette seconde conception. Il existe un réseau de symptômes, composé de symptômes interdépendants, utile à la pratique clinique. Ce que nous désignons comme une espèce psychiatrique correspond bien à cet ensemble (mécanistique) de symptômes. Ces symptômes constituent les conséquences « lisibles par le clinicien » de perturbations de mécanismes agrégés en un (autre) réseau, biologique (selon la conception 2). Plus précisément, nous soutenons que des mécanismes locaux (principalement biologiques) sont activés, c'est-à-dire que leurs interactions entraînent des symptômes chez un patient. Un mécanisme local correspond à un mécanisme relié à un symptôme et non à l'intégralité d'un trouble (comme le serait, par exemple, le « mécanisme rendant compte de la schizophrénie »). Trois éléments doivent être précisés au sujet de ce positionnement.

Premièrement, avec cette notion de mécanisme, nous introduisons dans notre thèse une position essentialiste. Cependant, nous prenons de la distance avec une position purement naturaliste, en affirmant que le trouble ne se définit pas au niveau de son mécanisme (biologique) – mais bien au niveau du réseau de symptômes en interaction. Certes, des mécanismes essentiellement biologiques rendent compte de symptômes, qui sont donc des effets d’entités non-visibles, situées « au-delà » de la phénoménologie clinique. Cependant, cette description ne tombe pas pour autant sous l’objection épiphénoméniste (c’est-à-dire le fait que les réseaux de symptômes ne sont que des réseaux de mécanismes produisant des symptômes), parce que la définition de ce qu’est un trouble psychiatrique ne se situe pas au niveau de ces mécanismes sous-jacents, mais bien au niveau des symptômes. Cette distinction est importante : tout l’intérêt de cette thèse est de caractériser un trouble psychiatrique pertinent pour le clinicien. Le recours au mécanisme n’est nécessaire que pour rendre compte de l’ancrage des symptômes dans des processus biologiques.

De plus, deuxièmement, nous marquons un écart important entre une « dysfonction se rapportant à l’ensemble des symptômes d’une catégorie » (comme l’avance par exemple la HDA) et un « mécanisme activé se rapportant à un symptôme préjudiciable » (comme nous l’avançons ici). Nous soutenons qu’un symptôme (S_a) est systématiquement relié à un mécanisme activé (M_a). Dans ce sens, nous ne réintroduisons pas de variable latente, qui sert à regrouper plusieurs symptômes en une seule entité sous-jacente. Un mécanisme témoigne d’un symptôme.

Troisièmement, comme vu dans la sous-section précédente, nous préférons soutenir la notion d’activation d’un mécanisme plutôt que celle de dysfonction. Bien entendu, nous pourrions considérer que la fonction est elle-même un simple mécanisme⁹⁸, mais cela ne changerait pas notre argumentation⁹⁹.

⁹⁸ Il existe en effet une littérature qui va dans ce sens (voir par exemple le chapitre « *Functional Mechanisms* » dans l’ouvrage de 2020 de Gualtiero Piccinini, *Neurocognitive Mechanisms*).

⁹⁹ Une telle conception ne devrait pas être trop rapidement rapprochée de la théorie du rôle causal de l’explication fonctionnelle de Cummins (1975, 1983). Celle-ci suggère que les fonctions permettent d’expliquer la contribution d’un certain nombre de composants d’une structure par rapport à un système plus large (contenant cette structure). Avec un autre vocabulaire, il s’agirait de dire, selon cette théorie, que la fonction d’un réseau de symptômes est de contribuer à l’environnement plus large dans lequel évolue ce réseau (ce patient). Cependant, une telle théorie nécessiterait de

Notre position est originale pour ces trois raisons : elle n'est pas anti-essentialiste, elle renvoie à l'activation d'un « mécanisme relié à un symptôme » plutôt qu'à un mécanisme relié à l'ensemble d'un trouble et elle s'appuie sur l'activation d'un mécanisme plutôt que sur une dysfonction.

4. Résumé

Nous soutenons qu'il n'y a pas de variable latente sous-jacente à un trouble psychiatrique et notamment qu'aucune dysfonction rendant compte de l'ensemble des symptômes d'un trouble psychiatrique n'est pertinente pour le clinicien. Au lieu de cela, nous proposons qu'une espèce psychiatrique soit définie comme un réseau de symptômes (préjudiciables – nous y reviendrons) soutenus par l'activation de mécanismes locaux. De plus, nous proposons que ce réseau soit constitué d'un noyau central qui renvoie à un prototype composé de symptômes qui sont en interactions mutuelles. Une espèce psychiatrique est constituée de l'ensemble des symptômes du noyau central d'un réseau de symptômes. En cherchant à dépasser la proposition nominaliste de Borsboom qui conçoit le symptôme comme une dénomination « flottante » dans la pratique clinique (Borsboom 2017), nous admettons que tout symptôme psychiatrique soit sous-tendu par l'activation d'un mécanisme local.

Ces deux propositions (noyau central et mécanisme local) sont développées et discutées dans la section suivante en relation avec la notion de seuil de préjudice. Nous y montrons la nécessité de reconsidérer le statut de préjudice tel qu'il est exprimé par le patient dans la pratique clinique.

mieux définir ce qu'est cette structure (ce réseau), cet environnement et cette contribution – ce que nous souhaitons éviter de faire dans ce travail.

Présumé II : Préjudice et analyse du symptôme préjudiciable

1. Rôle du préjudice en philosophie de la psychiatrie

À notre connaissance, le préjudice a pu être abordé sous deux formes en philosophie de la psychiatrie, soit comme conséquence d'une dysfonction, notamment au sein d'une analyse hybride comme la HDA (Wakefield 2013 ; Wakefield et Conrad 2019 ; Wakefield et First 2013), soit indépendamment de la notion de dysfonction. Elle a été définie comme une détresse ou un incapacité relatives aux normes, aux valeurs ou aux contextes (Cooper 2005).

Pour prendre l'exemple de la HDA, cependant, une confusion à propos de la notion de préjudice semblait présente dès l'article de 1992 de Jerome Wakefield. Celui-ci décrivait un trouble comme une dysfonction entraînant directement un préjudice. Il soulignait que le préjudice devrait être la conséquence d'une dysfonction, en écrivant : « Tout type de préjudice lié à la dysfonction d'un mécanisme interne pourrait être appelé un trouble ». Mais il ajoutait ce qui suit : « Notez que le lien, via les symptômes, avec le concept de préjudice est essentiel pour affirmer que l'échec du refoulement [la dysfonction] est un trouble ». Dans cet énoncé, le préjudice est « lié » à la dysfonction « via les symptômes ». Le préjudice est-il la conséquence d'une dysfonction, qui regroupe tous les symptômes d'une catégorie ? Ou bien le préjudice est-il lié à la dysfonction via un unique symptôme ? Dans la suite de cette section, nous développerons mieux la distinction entre un symptôme et un préjudice, mais nous pouvons d'ores et déjà assumer que le symptôme, en tant qu'élément recueilli en consultation par le clinicien, représente la « face clinique » de l'expérience vécue comme un préjudice par le patient.

Wakefield admet lui-même aujourd'hui que cette notion de préjudice « causé » par la dysfonction est ambiguë et reste une question irrésolue. Il écrit : « Il y a toujours eu une ambiguïté au sujet de l'expression 'causer de la détresse'. Dans quelle mesure cette causalité doit-elle être directe ? Est-ce qu'il est possible que si quelqu'un souffre du fait d'avoir un symptôme, la causalité est directe – ou le symptôme lui-même doit-il être une forme de préjudice ou causer « directement » du préjudice ? C'est une question encore ouverte à la controverse » (Wakefield, communication personnelle, 2023)¹⁰⁰. Ici, la question est de savoir si le symptôme est bien relié au préjudice. Or, précédemment, nous rapportons les termes de Wakefield affirmant que le préjudice était relié à un trouble dans son intégralité.

Cette ambiguïté se pose dans bien d'autres cas : par exemple, dans le DSM, le critère de significativité clinique indique que les symptômes doivent causer de la détresse ou une altération du rôle social. Cependant, quelle dysfonction est impliquée dans le préjudice tel qu'il est décrit au sein du « critère de significativité clinique » ? En quoi ce critère de significativité clinique (préjudiciable) est-il relié à une dysfonction ? Et à quel symptôme renvoie un préjudice relié à une dysfonction sous-jacente à un trouble ? En résumé, dans la HDA, il semble exister une confusion entre « la dysfonction reliée à / exprimée via une catégorie » (ou variable latente) et « la dysfonction reliée à / exprimée via un ensemble de symptômes ». Il existe donc un flou conceptuel concernant la nature du préjudice, que nous pourrions reformuler ainsi : si le préjudice est bien la conséquence d'une dysfonction (elle-même conséquence de l'altération d'un mécanisme supposé (Wakefield 1999)), la question de la relation des symptômes et des préjudices pose la question de savoir si l'ensemble des symptômes d'un patient sont la conséquence d'une unique dysfonction sous-jacente à un trouble ou si chaque symptôme, isolément, constitue lui-même une dysfonction préjudiciable.

Pour mieux comprendre l'origine de cette confusion, nous proposons de réancrer la HDA dans la pratique clinique. Sauf dans le cas des catégories particulières ne comportant qu'un seul symptôme (donc

¹⁰⁰ Il s'agit certes ici d'un cas particulier dans lequel le préjudice est lié à un symptôme, mais il n'est pas dû au symptôme lui-même (mais, par exemple, au fait même de se savoir présenter le symptôme). Mais cette subtilité n'est pas utile pour notre raisonnement.

monothétique), les patients, la société ou les cliniciens ne décrivent jamais de dysfonction (relative à l'évolution de leur espèce) qui serait liée à l'ensemble des symptômes qu'ils éprouvent ou recueillent. Autrement dit, les patients ne déclarent jamais qu'ils souffrent parce qu'ils ont une dysfonction indiquant « une dépression », mais plutôt qu'ils souffrent de leur anhédonie, de leur fatigue ou de leur aboulie. De même, personne dans la société ne dira : « Elle / il a un trouble appartenant au spectre de la schizophrénie, témoignant d'une dysfonction qui entraîne un préjudice », mais plutôt : « Son délire de mégalomanie et sa désorganisation sont chacun préjudiciables à leur manière ». Cette argumentation pourrait sembler purement rhétorique, car on pourrait soutenir que l'opinion des cliniciens ou des patients n'a pas d'importance dans la définition d'un trouble, que la plainte est une question sémantique et qu'un patient atteint de cancer, qui ne sait rien de sa métastase osseuse, se plaindra de sa douleur sans que la plainte invalide les connaissances médicales sur la tumeur primitive. Or, ce sont précisément ces arguments que nous cherchons à critiquer en proposant une définition théorico-clinique, supposée intuitive pour le clinicien. En d'autres termes, nous proposons que ce que le clinicien rencontre devrait *prima facie* constituer la base de la définition d'un trouble. Cela nous fait rejeter ce contre-argument minimisant l'avis du clinicien ou du patient dans la définition du trouble¹⁰¹. D'ailleurs, la plupart des débats sur le normal et le pathologique sont rattachés à un « sens commun » qui permet de guider les orientations conceptuelles (par exemple, il serait délicat de soutenir une définition du trouble psychiatrique qui considère que l'homosexualité est un trouble psychiatrique)¹⁰². Pour donner un autre exemple de ce sens

¹⁰¹ Rappelons que notre intérêt est de comprendre le trouble psychiatrique du point de vue du clinicien et, par extension, du point de vue de celui qui en fait l'expérience. Sur le plan philosophique, comme nous le verrons plus loin, les symptômes sont cependant sous-tendus par des mécanismes. Ainsi, en plus d'être utile en pratique clinique, l'approche par le symptôme peut également s'avérer pertinente pour des fonctions de recherche ou de pédagogie (entre autres), puisqu'il sera intéressant d'expliquer (en recherche et / ou en pédagogie) que ces symptômes sont reliés à des métastases osseuses.

¹⁰² Nous n'entrerons pas dans les considérables débats portant sur cette question du sens commun, abordé en philosophie analytique, qui ne sont pas nécessaires pour notre thèse. Cependant, comme avec cet exemple de l'homosexualité, il est important de préciser que ce que nous désignons par « sens commun » comporte nécessairement une dimension culturelle et reste contingente aux facteurs contextuels propres à une temporalité et une spatialité données.

commun, dans l'HDA au moins, la notion de « preuve circonstancielle » est centrale¹⁰³. Sur cette base du sens commun, les symptômes d'un patient ne peuvent que difficilement correspondre, *prima facie*, à la conséquence d'une unique dysfonction sous-jacente à un trouble. En d'autres termes, il semble difficile d'aisément et communément postuler une unique variable latente permettant d'expliquer l'agrégation des différents symptômes d'un trouble psychiatrique. Comment concevoir qu'il existerait une dysfonction unique à l'origine de la tristesse de l'humeur, de l'aboulie, du ralentissement psychomoteur et de l'insomnie (sans compter que l'hypersomnie peut aussi compter comme un critère de l'épisode dépressif caractérisé). Wakefield propose une réponse apparemment claire basée sur une théorie évolutionniste : le mécanisme dysfonctionnel à l'origine du regroupement de symptômes a été sélectionné naturellement – une réponse qui suscite néanmoins des critiques¹⁰⁴.

Nous proposons de mieux explorer ce qui est « contenu » au sein d'une catégorie (du DSM), réunie par une dysfonction – réflexion qui peut s'étendre aisément aux fonctions psychologiques des RDoC ou aux dimensions réunies par des construit statistiques au sein de la HiTOP. Les symptômes et les préjudices sont exprimés par le patient dans la pratique clinique et recueillis par le clinicien.

Nous considérons qu'un symptôme peut constituer une dysfonction préjudiciable, ce qui revient à dire que chaque dysfonction et chaque préjudice sont « intégrés » à un symptôme. Qu'est-ce que cela signifie d'affirmer qu'un préjudice est « intégré » au sein d'un symptôme (Cooper 2005, 2013a, 2014, 2020) ?

¹⁰³ La preuve circonstancielle est un terme issu du vocabulaire juridique permettant de désigner une preuve indirecte qui permet de guider l'action. Cet argument est utilisé à de multiples reprises par Wakefield depuis plus de 30 ans. On retrouve par exemple dès son article 1992 : « Les jugements sur ce qu'est un trouble peuvent être basés sur des preuves circonstancielle lorsque la connaissance des mécanismes fait défaut, comme lorsque nous déduisons que la cécité et les hallucinations sont des troubles, sans rien comprendre au fonctionnement de la perception ».

¹⁰⁴ Bien que ce point soit hors de portée de notre argumentation car il n'est pas utile au développement de celle-ci, il faut noter que les théories évolutionnistes ne sont pas incompatibles avec notre définition d'une espèce psychiatrique fondée sur les réseaux de symptômes. En effet, une des hypothèses permettant d'expliquer ce qui rassemble tel ou tel symptôme au sein d'un réseau peut s'expliquer par la pression de sélection évolutive (voir notamment les arguments de Nesse, 2023, dans *World Psychiatry*). Cependant, étant donné que nous cherchons ici à mieux cerner une définition utile et intuitive pour le clinicien, ces origines évolutionnistes, conceptuellement compatibles, ne nous intéressent pas et ne seront pas plus discutées.

Comme l'avance Wakefield lui-même, « les manifestations cliniques sont simplement la face clinique [*the clinical side*] du préjudice ; certaines sont des indicateurs de dysfonctions et d'autres sont à la fois des indicateurs de dysfonction et de préjudice. Par exemple, dans la dépression majeure, un profond sentiment d'inutilité et un retard psychomoteur sont à la fois préjudiciable à première vue, mais leur nature inhabituelle et dramatiquement inadaptée suggère aussi une dysfonction interne des systèmes affectifs » (Wakefield, communication personnelle, 2023). Ainsi, un symptôme peut être l'indicateur clinique d'un préjudice, mais également l'indicateur clinique d'une dysfonction¹⁰⁵. De fait, en tant qu'un symptôme peut constituer l'indicateur clinique d'un préjudice (ou la « face cliniquement visible du préjudice »), par la suite, nous discuterons de l'implication de tels « symptômes préjudiciables »¹⁰⁶.

2. Notion de symptôme préjudiciable

Nous soutenons qu'une conception philosophique fondée sur des « symptômes préjudiciables » est plus cohérente avec la position implicite du clinicien qu'une conception fondée sur la notion de dysfonction préjudiciable, telle qu'elle peut être présentée dans la HDA. Trois points doivent être précisés, concernant : la manière dont le préjudice est relié au symptôme ; le fait que certains symptômes seulement sont préjudiciables ; la définition de ce qu'est un préjudice pour la philosophie de la médecine.

Si un préjudice est difficilement « lié » à un ensemble de symptômes formant une catégorie dans le DSM, comment un préjudice serait-il mieux « lié » à un unique symptôme ? Plutôt que de postuler, en pratique

¹⁰⁵ On retrouve ici la notion du symptôme comme « indice » d'une cause sous-jacente (Demazeux, 2019).

¹⁰⁶ Nous ne discuterons plus, par la suite, des « symptômes dysfonctionnels », qui ne sont pas spécialement utiles pour notre argumentation. Mais il faut retenir qu'il existe en effet certains symptômes qui représentent la face cliniquement visible de dysfonctions utiles à considérer pour définir un trouble et d'autres qui ne représentent cliniquement pas de dysfonction. Par exemple, le refus d'alimentation est un symptôme de l'anorexie mentale mais ne correspond pas à la dysfonction d'un mécanisme orexigène – mais plutôt cognitif ou émotionnel. Wakefield écrivait : « La nature précise des symptômes qui résultent [du refoulement manqué, constituant un trouble] dépend de nombreux autres facteurs idiosyncrasiques [que le refoulement manqué lui-même] » (Communication personnelle, 2023). Berrios a largement détaillé ce point en insistant sur la notion de « soupe primordiale » et Kirmayer y a intégré une perspective écosociale.

clinique, l'existence d'un lien de causalité entre un préjudice et un symptôme, nous proposons que le préjudice soit une des composantes intrinsèques de l'un de ces symptômes.

Plus précisément, comme évoqué plus haut, le préjudice et le symptômes constituent les deux faces d'une activation de mécanisme : le préjudice est le vécu phénoménologique du patient (ou le vécu préjudiciable d'un individu tel qu'il est décrit selon les valeurs sociales) et le symptôme en est la dénomination métaphorique émanant de ce préjudice et lisible par le clinicien (au prisme de son bagage théorique, de sa formation et de son expérience – et tout l'enjeu d'un « appareil sémiologique » pourrait être discuté ici (Demazeux, 2019)). Le symptôme est la face clinique de l'expérience vécue comme un préjudice par le patient.

Cette définition d'un symptôme préjudiciable est renforcée par le fait que la présence dans le DSM du critère de signification clinique (stipulant que certains symptômes « doivent causer des souffrances ou des répercussions importantes ») est présent dans la majorité des troubles où les symptômes eux-mêmes ne causent aucun préjudice évident (Wakefield 2013 ; Wakefield et Conrad 2019 ; Wakefield et First 2013). Autrement dit, quand le préjudice ne semble pas « intuitivement » intégré à un symptôme, il est ajouté dans un critère supplémentaire pour certaines catégories du DSM. Plus empiriquement encore, pour donner un autre exemple soutenant la notion de symptôme préjudiciable, l'idée que tout symptôme peut être considéré comme préjudiciable (selon le contexte, le patient, etc.) a été reprise dans les enquêtes épidémiologiques et notamment dans l'étude *Epidemiological Catchment Area*, la première grande enquête épidémiologique psychiatrique basée sur le DSM aux États-Unis (Eaton et al. 1981 ; Shapiro et al. 1984) ou dans la constitution d'une cohorte étudiée actuellement comme la *National Epidemiologic Survey on Alcohol and Related Conditions* (NESARC). Dans ces enquêtes, lorsqu'un symptôme était rapporté par un patient, l'enquêteur posait une série de questions sur sa gravité et le symptôme n'était enregistré qu'à partir d'un certain seuil¹⁰⁷.

¹⁰⁷ Ce point nous avait été proposé par Jerome Wakefield lui-même (communication personnelle, 2021).

Tous les symptômes préjudiciables du noyau central d'une condition psychiatrique n'ont pas la même valeur clinique en termes de préjudice pour un patient (Tabb 2019 ; Zachar 2021). Certains d'entre eux, à certains moments, dans certains milieux et chez certains patients, sont préjudiciables. Il peut en effet y avoir des symptômes qui ne causent aucun préjudice. Par exemple, les personnes qui entendent des voix peuvent dire que leurs hallucinations acoustico-verbales contribuent à leur bien-être et des individus présentant des difficultés dans les interactions sociales peuvent affirmer ne pas ressentir de préjudice concernant ce type de communication qualitativement différente engagé avec autrui.

Il nous faut également définir ce que nous nommons un préjudice. Contrairement à la déficience (par exemple, l'absence congénitale de membres) ou au handicap (qui fait référence à une incapacité à accomplir les activités quotidiennes de la vie (Oliver, 1996)), le préjudice a été décrit comme la détresse ou l'incapacité vécue par une personne ou reflétée dans les valeurs d'une société (Wakefield, 1992). Les préjudices surviennent relativement aux normes culturelles acceptées dans un intervalle de temps et une culture donnés. Dans le présent chapitre, nous nous appuyons principalement sur la description de ce qu'est un préjudice telle qu'elle a été proposée par Wakefield « conformément aux valeurs sociales » (Wakefield, 1992). Nous définissons le préjudice comme tout symptôme décrit par un individu (et recueilli par un clinicien) qui « compromet la capacité d'une personne à mener une vie normale ou sans entrave selon les normes » d'un espace donné (par exemple, culturel ou communautaire) (Cooper, 2021)¹⁰⁸.

¹⁰⁸ Nous justifions notre utilisation de cette définition par le fait que la notion de préjudice est débattue en philosophie de la psychiatrie depuis plus de trente ans et que malgré ses limites, ses écueils et ses critiques (De Block et Sholl 2021 ; Powell et Scarffe 2019a, 2019b), la définition de Wakefield du préjudice relative au seuil et à l'altération de la vie selon les valeurs sociales et culturelles est fondamentale pour notre argumentation. Certes, il pourrait être intéressant de chercher à démontrer à *nouveau* pourquoi le seuil de préjudice décrit par Wakefield (Spitzer et Wakefield 1999 ; Wakefield et Spitzer 2002) ou par Spitzer (Spitzer 1999) est intéressant pour la pratique clinique, mais nous jugeons que ces discussions ont déjà été largement abordées dans la littérature (Faucher et Forest 2021). Le détail de ces débats ne nous intéressent finalement pas pour notre argumentation d'une espèce psychiatrique fondée sur une approche par le symptôme.

La définition du préjudice peut également être fournie sur la base de sa relation au mécanisme qui le sous-tend. En ce sens, un symptôme non préjudiciable (comme une simple fluctuation de l'humeur) rend compte d'un mécanisme inactivé ou activé en dessous d'un seuil. Une symptomatologie préjudiciable est la conséquence d'un mécanisme activé au-dessus d'un certain seuil.

Enfin, il s'agit de donner des précisions sur les éléments qui entrent en interaction dans le réseau : est-ce les symptômes, les préjudices, les mécanismes sous-jacents ? Prenons l'exemple d'un délire de persécution survenant à la suite d'une hallucination acoustico-verbale. En retour, le délire aggravera l'hallucination, les deux symptômes étant fortement reliés de manière bidirectionnelle. Une telle relation peut être décrite de trois manières : i) en termes sémiologiques, de manière littérale¹⁰⁹, la description clinique « délire de persécution » sera considérée comme augmentant la description clinique qu'est « l'hallucination » ; ii) en termes normativistes, le préjudice lié au délire aggravera fortement le préjudice lié à l'hallucination ; iii) d'un point de vue mécanistique, le mécanisme sous-jacent au délire activera le mécanisme sous-jacent à l'hallucination¹¹⁰. Ces trois descriptions ne s'excluent pas et sont toutes trois valables à trois niveaux d'explication différents, afin de définir ce qu'est une espèce psychiatrique pour le clinicien. À partir de là, nous pouvons proposer que le symptôme soit considéré comme la description clinique (i) d'un préjudice (ii) relié à un mécanisme activé (iii).

3. Préjudices et espèce psychiatrique

Qu'en est-il de la possibilité d'utiliser des symptômes préjudiciables en interactions mutuelles pour définir des espèces psychiatriques cliniquement utiles ? Nous avons dit que certains symptômes

¹⁰⁹ Nous employons le terme de « littéral » et avons précédemment employé les termes de « stipulatif » ou « ostensif ». Ces termes renvoient au fait que les symptômes sont avant tout des « dénominations » et que nous ne nous prononçons pas sur leur nature ou leur fonction.

¹¹⁰ Comme nous l'avons décrit plus haut, l'existence d'un mécanisme – ou même d'un réseau de mécanismes – sous-jacent n'entre pas dans la définition du trouble et n'a d'ailleurs pas de valeur directement pertinente pour une fonction clinique telle que nous l'avons défini. En ce sens, c'est le réseau de symptômes qui nous intéresse dans cette argumentation, notamment car c'est lui qui est pertinent.

pouvaient être présents chez un patient, mais non préjudiciables, comme une humeur légèrement déprimée ou une anhédonie légèrement marquée. Un patient peut même signaler la présence de l'ensemble des neuf symptômes appartenant au construit « trouble dépressif caractérisé », sans qu'aucun d'entre eux ne dépasse le seuil de préjudice historiquement considéré – et socialement admis – comme significatif. Il pourrait donc exister un ensemble de symptômes préjudiciables chez un patient sans que celui-ci ne soit considéré comme porteur d'un trouble psychiatrique.

Pour éclairer cette question, nous soutenons qu'une espèce psychiatrique doit présenter : i) au moins un préjudice s'exprimant cliniquement sous la forme d'un symptôme ; ii) *et* que ce ou ces symptômes préjudiciables doi(ven)t appartenir au noyau central prototypique des affections considérées par la discipline psychiatrique ; iii) *et* que ce préjudice ou la somme des préjudices doit dépasser le seuil de signification clinique communément admis sur la base des valeurs de la société ; iv) *et* au moins un des préjudices doit être le reflet d'une activation d'un mécanisme (biologique).

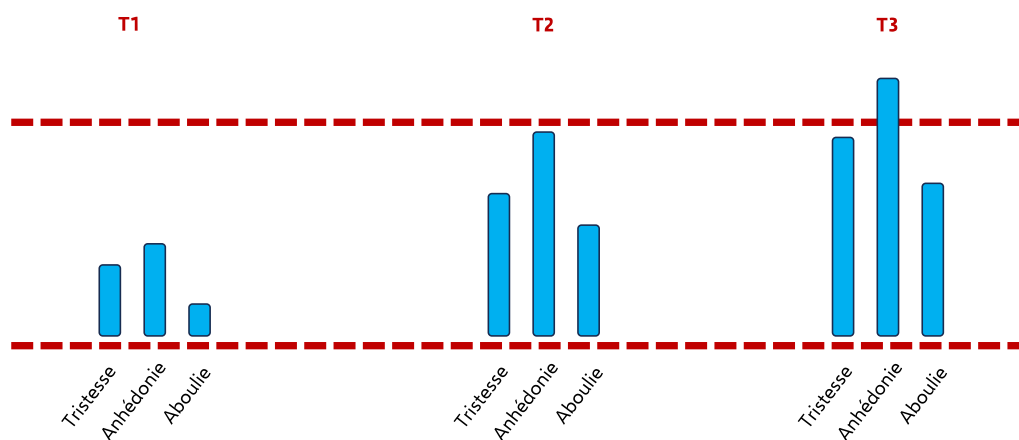


Figure 6 – Seuil de préjudice considéré comme pathologique selon les valeurs de la société. Si au moins un symptôme préjudiciable franchit le seuil, le patient sera considéré comme ayant un trouble psychiatrique. S'il s'agit d'un préjudice qui n'est ni intégré, ni relié directement à un symptôme sous-tendu par un mécanisme activé, il ne s'agit pas d'une espèce psychiatrique. T1, T2, T3 : trois temporalités arbitraires. Le nom des symptômes dans la Figure sont donnés à titre indicatif. Nous verrons ci-dessous qu'une condition supplémentaire d'actionnabilité est nécessaire pour qu'une condition soit considérée comme un trouble.

Concernant (iii), pour être désigné comme un trouble psychiatrique, il faut qu'au moins un préjudice franchisse un seuil considéré comme pathologique selon les valeurs de la société (**Figure 6**). Cependant, l'une des difficultés de la psychiatrie d'avant les années 1980 était d'éviter de rendre pathologique un symptôme préjudiciable isolé. C'est d'ailleurs une des raisons qui a motivé certains membres de l'École de Saint-Louis à se pencher sur la question du diagnostic et à concevoir une approche par la catégorie (par exemple, en considérant que la tristesse isolée ne devait pas être pathologique tandis que l'épisode dépressif caractérisé était pathologique). Nous verrons tout de suite comment éviter de retomber dans cette difficulté en détaillant la notion de réseau figé, c'est-à-dire de « rigidité » entre les symptômes.

Concernant (iv), comme proposé précédemment, la définition d'une espèce psychiatrique utile à la pratique clinique nécessite de considérer l'activation d'au moins un mécanisme biologique sous-jacent à un symptôme préjudiciable. En l'absence d'activation d'au moins un mécanisme, tout préjudice non médical (n'appartenant pas à la psychiatrie) relèverait du champ médical (par exemple, un préjudice lié à la précarité ou à l'évasion fiscale)¹¹¹ (Tsou 2021a). La perte d'un proche, la découverte d'un cancer, l'échec d'un examen, ou toutes autres sortes de raisons, parfois passagères, constituent des souffrances psychologiques bien identifiées. Ces conditions causent sans aucun doute un préjudice à la personne. Faut-il considérer comme des ensembles de symptômes préjudiciables ces états plus ou moins transitoires d'angoisse, de tristesse ou d'inquiétude ?

Afin de considérer médicalement qu'une condition soit une espèce psychiatrique, il faut qu'au moins un préjudice, appartenant au noyau central des symptômes (par exemple, l'exclusion sociale), soit : i) soit intégré directement à un symptôme (par exemple, l'exclusion sociale est un préjudice intégré à un « symptôme négatif » retrouvé classiquement un tableau constitué d'autres symptômes psychotiques) ; ii) soit une conséquence directe d'un symptôme (par exemple, l'exclusion sociale est un préjudice

¹¹¹ Puisque nous ne cherchons pas à délimiter les conditions d'accès aux soins, nous ne voulons pas dire ici que les cliniciens ne devraient pas s'occuper de personnes dans de telles situations, comme s'ils pratiquaient une clinique « pure », exempte d'autres épaisseurs appartenant aux champs sociaux et conjoncturels. Plutôt, de manière minimale, nous cherchons à proposer une définition d'une espèce psychiatrique pertinente pour la clinique telle qu'elle est pratiquée en face-à-face.

survenant par conséquence directe du symptôme qu'est le délire paranoïaque). Cette présence minimale d'un préjudice intégré ou conséquence directe d'un symptôme n'exclut pas l'existence de préjudices qui seraient les conséquences d'environnements socialement préjudiciables (par exemple, l'exclusion sociale faisant suite à la discrimination à l'encontre des personnes souffrant de délires paranoïaques)¹¹². Nous proposons donc de différencier trois types de préjudices sur le plan épistémologique : ceux qui sont « intégrés » à un symptôme et conséquence directe d'un mécanisme activé ; ceux qui sont les conséquences directes d'un symptôme ; ceux qui sont les conséquences d'environnement socialement préjudiciables. Seuls les deux premiers semblent valables pour considérer une condition comme une espèce psychiatrique.

4. Réseau figé

Nous avons dit que les espèces psychiatriques pouvaient être définies comme le noyau central d'un ensemble de symptômes préjudiciables soutenus par un ensemble de mécanismes activés. Nous allons voir en quoi notre définition d'une espèce psychiatrique est pragmatique, au sens d'Engelhardt ou de Canguilhem, c'est-à-dire qu'elle appelle à l'action. Nous allons voir plus précisément en quel sens il est nécessaire d'intervenir afin de restaurer l'état initial du réseau.

Nous venons de voir qu'il fallait au moins un préjudice franchissant un seuil considéré comme pathologique selon les valeurs de la société parle de trouble psychiatrique (**Figure 6**). Cependant, afin d'éviter que tout symptôme préjudiciable isolé soit considéré comme pathologique, nous introduisons le fait qu'un ensemble de symptômes ne devait être considéré comme pathologique que lorsque cet

¹¹² Ce point résout la question précédente, posée par Wakefield, de savoir si le fait de souffrir de se savoir présenter un symptôme constitue un préjudice : oui, cela constitue un préjudice, mais ce préjudice ne « compte pas » si au moins un autre préjudice relié à un mécanisme activé, via un symptôme du noyau central, n'est pas également présent.

ensemble est rigide, figé¹¹³ (Borsboom et al., 2019). Ce critère supplémentaire, original, nous semble nécessaire dans la définition d'une espèce psychiatrique utile à la pratique clinique.

Cette notion de rigidité nous permet de mieux différencier une condition non pathologique d'un trouble psychiatrique. Nous définissons l'absence de trouble psychiatrique comme l'état stable d'un réseau faiblement connecté. Les connexions entre les symptômes correspondent aux forces d'associations, qui se comprennent aisément dans la pratique clinique : deux symptômes sont fortement associés s'ils sont tous deux reliés avec une forte intensité. Par exemple, la peur des souris (symptôme 1) peut entraîner un évitement comportemental (symptôme 2) de certains lieux, comme les greniers ou les ruelles. Ces deux symptômes sont considérés comme étant fortement reliés. Cependant, si un patient éprouve une peur (même intense) mais n'a pas d'évitement à chaque fois qu'il rencontre une souris, ces deux symptômes sont considérés comme étant faiblement associés. De telles connexions faibles sont le reflet de la capacité d'adaptation du patient aux changements de son environnement. En d'autres termes, même si des symptômes sont potentiellement isolément particulièrement préjudiciables, si la connexion entre eux est faible, l'individu saura s'adapter et moduler ses comportements, émotions et cognitions. De fait, le réseau de symptôme ne sera pas considéré comme figé, rigide. Le patient ne sera donc pas diagnostiqué comme porteur d'un trouble psychiatrique. Ainsi, la condition de « rigidité » d'un réseau est particulièrement importante pour ne pas diagnostiquer tout symptôme intense comme une trouble psychiatrique.

Les réseaux non pathologiques peuvent s'adapter et changer de manière flexible en réponse aux perturbations environnementales (Herrman et al. 2011). Un réseau faiblement connecté est donc dit « robuste », dans le sens où la robustesse permet une adaptation aux perturbations (Kitano 2004 ; Lesne 2008, p. 200 ; Stelling et al. 2004). Il est alors dit sain, non pathologique. L'adaptation commune des symptômes facilite l'adaptation sociale d'un individu, même s'il souffre de préjudices liés à des mécanismes activés. Comme ces préjudices restent en dessous du seuil considéré par les valeurs de la

¹¹³ Cette rigidité renvoie à la notion d'hystérésis en sciences des systèmes dynamiques.

société, le réseau de symptômes peut se modifier en fonction des nécessités liées aux changements contextuels et environnementaux.

Une espèce psychiatrique « complexe » correspond, à l'inverse, à une difficulté d'adaptation (ou de contrôle) lors de laquelle le réseau, figé, n'arrive plus à s'adapter aux perturbations sans intervention thérapeutique. C'est donc bien la notion de besoin d'intervention par la société qui entre dans la définition d'une telle espèce psychiatrique. Par exemple, dans le cadre de la pédophilie ou des troubles antisociaux, les individus subissent un préjudice en raison d'un manque de contrôle de leurs symptômes, notamment parce que leurs émotions (qui sont des symptômes préjudiciables) s'influencent fortement les uns les autres (par exemple, (Gerin et al. 2017)). Ce réseau d'influences est tellement important que cette rigidité ne donne plus accès à – par exemple – un rétablissement par des méthodes de psychologie populaire ou à la métacognition du sujet lui-même. Le patient qui décrit ce réseau figé d'influences requiert une intervention médicale. Il est crucial de noter que cette intervention doit provenir spécifiquement de la médecine et non de l'individu lui-même ou de la société. En effet, une espèce psychiatrique ne peut être qualifiée en tant que tel que lorsque l'autonomie de la personne est altérée au point de l'empêcher s'adapter à un contexte social donné. En ciblant le préjudice – défini comme une détresse ou un incapacité relatives aux normes, aux valeurs ou aux contextes –, la médecine contribuerait ainsi à favoriser l'adaptation du patient.

—

Dans la section suivante, nous résumons les implications de l'identification et de la déconstruction de ces deux présupposés (portant sur les notions d'activation d'un mécanisme et de préjudice) et proposons une définition potentiellement pertinente de ce qu'est une espèce psychiatrique pour la pratique clinique, nommée espèce complexe. C'est en proposant une telle définition que nous pourrions comprendre comment une telle espèce psychiatrique peut être utile pour le clinicien. À travers cette définition, nous

pourrons ainsi soutenir une approche par le symptôme utile pour répondre avant tout aux nécessaires fonctions cliniques de la discipline.

Une définition potentiellement pertinente de ce qu'est une espèce psychiatrique pour la pratique clinique

Dans les sections précédentes, plutôt que d'adopter une approche par la catégorie / dimension, dans laquelle les espèces psychiatriques sont définies sur la base d'une variable latente (Cooper, 2002 ; Wachbroit, 1994) et principalement en fonction d'espèces diagnostiques dérivées du DSM, nous avons soutenu que les espèces psychiatriques pouvaient être comprises en termes de symptômes en interactions mutuelles (Borsboom 2017 ; Tsou 2016).

Nous avons argumenté que certains des symptômes du noyau central prototypique de ce réseau constituent la face clinique des préjudices décrits par un patient ou considérés comme tels par les valeurs de la société (Cooper 2005). De plus, nous avons avancé que les symptômes préjudiciables sont soutenus par l'activation d'un mécanisme local. Enfin, nous avons montré l'importance de considérer un ensemble de symptômes figés dans des interactions mutuelles.

1. Avantage de considérer des espèces psychiatriques utiles pour le clinicien

Une approche cliniquement pertinente fondée sur les symptômes présente des avantages cliniques et possède également des répercussions et engagements au niveau institutionnel, scientifique et psychométrique.

Cliniquement, cette approche est intuitive pour le clinicien dans sa pratique clinique, c'est-à-dire qu'elle lui permet d'accéder à des heuristiques diagnostiques, prédictives, pronostiques et thérapeutiques pertinentes et rapides. Comme nous l'avons dit en introduction avec l'exemple des thérapies cognitivo-comportementales, un nombre croissant de recherches soutient que le raisonnement clinique est largement guidé par une conception en réseau, qui construit des réseaux (cognitifs) sur la base de variables en interaction (Bhugra et al. 2011). Par exemple, le clinicien recherche des hallucinations acoustico-verbales pour tenter d'identifier si celles-ci alimentent un délire et il s'attache à définir si c'est ce délire qui conduit au retrait social. Sa question sera de savoir si c'est le délire ou plutôt une anhédonie qui est lié(e) aux hallucinations constituées d'injonctions suicidaires, conduisant elles-mêmes à un tel repli social. Le DSM lui-même est en partie construit, implicitement, sur de telles influences, par exemple dans le cadre de l'état de stress post-traumatique (le critère B.2. est celui-ci : « Rêves répétitifs provoquant un sentiment de détresse dans lesquels le contenu et / ou l'affect du rêve sont liés à l'événement / aux événements traumatiques »). Enfin, la notion de noyau central, constitué de tels symptômes en interactions mutuelles, évite de fixer une limite claire entre des catégories (comme cela tente d'être fait dans les approches par la catégorie / dimension). Ce flou périphérique (réduisant les « zones de rareté » – Kendell, 1968) est particulièrement intuitif pour le clinicien qui ne cherche pas à tracer des « lignes dans le sable » délimitant un trouble d'un autre (Kendell et Jablensky, 2003).

Sur le plan institutionnel, cette approche par le symptôme n'implique pas de modifier le seuil diagnostique traditionnellement proposé par la HDA (et le DSM). En effet, pour un patient donné, un même seuil sera fixé, que l'on utilise la définition générique de la HDA ou la définition théorico-clinique. Avec les deux définitions (HDA et théorico-clinique), le préjudice est lié respectivement à une dysfonction ou à un mécanisme, évitant ainsi une approche purement normative et ses conséquences,

comme la surmédicalisation. Cette dernière est également limitée par l'existence d'un seuil de préjudice, permettant de ne pas médicaliser des états simplement indésirables en dessous de ce seuil (par exemple, le chagrin).

Sur le plan scientifique, l'approche par le symptôme fait écho aux défis posés par la médecine de précision. Celle-ci a plus d'espoir d'identifier des mécanismes physiopathologiques au niveau des symptômes qu'au niveau des troubles dans leur globalité (Fernandes et al. 2017 ; Gómez-Carrillo et al. 2018) (**Annexe 4 – « Le spectre d'essentialité »**). Par exemple, les programmes scientifiques de psychiatrie computationnelle s'orientent vers l'identification du mécanisme de l'hallucination que du mécanisme de la catégorie de schizophrénie (**Annexe 4 – « Le spectre d'essentialité »**). De plus, le noyau central permet de garantir une stabilité aux espèces psychiatriques pour la recherche, tout en laissant la périphérie du réseau évoluer au gré des découvertes scientifiques. Par exemple, des biomarqueurs pourraient être ajoutés dans la périphérie du réseau de la schizophrénie au fil des progrès de la science, si ces biomarqueurs influencent les interactions entre symptômes, mais sans qu'ils ne modifient la définition de ce qu'est une telle espèce psychiatrique.

Enfin, d'un point de vue psychométrique, cette approche et sa définition cliniquement utiles s'inscrivent dans une logique de raffinement de la validité de contenu, notion largement écartée des débats depuis l'article de 1955 de Cronbach et Meehl (Cronbach et Meehl 1955). Interroger la validité de contenu pourrait être considéré comme une tentative de réponse à Sadowsky, qui écrivait que « le menu des symptômes du DSM forme une constellation de signes de souffrance qui ont été regroupés dans des contextes disparates » (Sadowsky, 2021). Cette constellation de symptômes préjudiciables devrait être mieux explorée en les structurant au sein des réseaux. De tels réseaux, en présentant certains symptômes plus liés que d'autres et formant un noyau central prototypique, réinterrogent sous un autre angle la nature des éléments et des relations qui constituent une espèce psychiatrique. En pouvant s'appuyer sur les études empiriques identifiant les symptômes centraux de la psychopathologie et en ouvrant la

discussion sur la validité de contenu, nous espérons qu'une telle définition relancera dans une nouvelle direction les discussions en philosophie de la psychiatrie.

2. Proposition de définition d'une espèce psychiatrique

Sur la base de la déconstruction de ces deux présupposés et de cette dernière analyse concernant l'activation des réseaux, nous proposons qu'une espèce psychiatrique utile à la pratique clinique puisse être conçue comme « un ensemble de symptômes appartenant au noyau central prototypique de la psychiatrie, organisés en réseau, impliquant au moins un préjudice et dont l'ensemble des préjudices dépasse le seuil de signification socialement accepté selon les valeurs sociales admises dans une société pour une condition donnée, chacun de ces symptômes étant lié à un mécanisme activé. Un tel réseau de symptômes est figé dans un état qui ne permet pas la diminution spontanée du préjudice relié à ces symptômes sans intervention thérapeutique ». Pour résumer de manière pédagogique cette définition, les troubles psychiatriques peuvent être considérés comme des ensembles de symptômes interconnectés préjudiciables pour le patient. Ils ne peuvent être modifiés que par une intervention de la médecine. Il est indéniable qu'inclure la médecine elle-même dans la définition du trouble peut sembler audacieux. Cependant, il semble important de montrer qu'une espèce psychiatrique est nécessairement définie comme telle dès lors que les soins menés par la personne elle-même, ses proches, la société – et même d'autres professionnels de santé comme les psychologues, même si cette affirmation pourrait être discutée – sont devenus insuffisants pour répondre à l'appel à l'aide et / ou à l'absence d'autonomie du sujet incapable de s'adapter dans un contexte social donné. La médecine aiderait alors à l'adaptation du sujet.

Nous donnons ci-dessous cinq exemples d'utilisation de cette définition pour le clinicien. Premièrement, un patient présente un schéma de comportements alimentaires restrictifs (jeûne et hyperactivité

physique), qui le conduit chaque jour à avoir des épisodes d'hyperphagie (apport alimentaire excessif), entraînant un sentiment de culpabilité par peur de prendre du poids, ce qui entraîne en retour des comportements alimentaires restrictifs. Dans ce réseau de symptômes, il existe des nœuds (par exemple, la culpabilité) que sont les symptômes potentiellement préjudiciables et des connexions entre les nœuds que représentent les interactions entre ces symptômes. Ce premier exemple illustre la notion de dépendance (dite « causale » – voir (Gauld, 2021) sur la notion de causalité dans les réseaux de symptômes) entre les symptômes.

Deuxièmement, une patiente apparemment en bonne santé se présente chez un psychiatre, sur les conseils de sa famille, pour demander de l'aide au sujet d'hallucinations acoustico-verbales. Celles-ci lui donnent l'impression d'être surveillée. Cependant, ces hallucinations acoustico-verbales ne nuisent pas à la patiente dans son quotidien et n'affectent pas son entourage ni sa communauté. Dans un contexte professionnel, rien ne laisse transparaître qu'elle puisse éprouver de tels symptômes et son travail n'en est pas affecté. Elle considère d'ailleurs sa condition et son vécu comme une forme de « neuroatypicité ». La patiente n'a pas de trouble psychiatrique selon la définition théorico-clinique précédente, car le seuil de préjudice reste bas, malgré la présence et la potentielle influence mutuelle des symptômes. Ce deuxième exemple illustre la notion de seuil de préjudice décrite dans la définition ci-dessus.

Troisièmement, un patient décrit des difficultés liées à son travail, qui n'ont pas d'impact financier sur sa vie quotidienne ou celle de sa famille compte tenu de ses ressources et de ses économies élevées. Il présente un préjudice professionnel important lié à la perception d'une injustice due à un poste qu'il estimait mériter mais qui a été confié à un collègue. Dans ce cas, ce que rapporte le patient n'a d'impact ni sur son humeur (niveau d'anxiété), ni sur sa vie quotidienne. Il existe donc des préjudices chez ce patient, mais aucun symptôme relié à un mécanisme (biologique) activé ne peut être décrit ni retrouvé.

Il n'y a donc pas de trouble psychiatrique chez ce patient¹¹⁴. Ce troisième exemple illustre la notion de préjudice directement lié aux symptômes, proposée dans la définition ci-dessus.

Quatrièmement, un patient présentant un état de manque, décrit par le symptôme de « craving », une envie ou besoin irréprensible de consommer une substance, qui constitue un symptôme central parmi d'autres symptômes liés à l'usage de substance (Drummond et al. 2000 ; Sayette 2016 ; Serre et al. 2018 ; Vafaie et Kober 2022 ; Gauld et al., 2023). Le craving est sous-tendu par un mécanisme physiopathologique identifié entre autres au niveau des circuits de la dopamine et de la sérotonine des gyri para-hippocampiques et fusiformes, du putamen, du cortex cingulaire antérieur, de l'amygdale et du cortex orbitofrontal (Robinson et Berridge 1993). Même s'il existe d'autres mécanismes à d'autres échelles du vivant et d'autres niveaux de compréhension, ce qui nous importe est qu'il existe *au moins un* mécanisme biologique. Si le craving était la seule manifestation des comportements de la vie quotidienne, sans être lié à au moins un mécanisme, il n'aurait pu être intégré dans la définition précédente. Ce quatrième exemple illustre l'importance de considérer les symptômes comme directement liés à un mécanisme activé.

Cinquièmement, dans le contexte d'un patient présentant des symptômes de stress post-traumatique, le traitement par blocage β -adrénergique lors de séances de désensibilisation et de retraitement des mouvements oculaires (*Eye Movement Desensitization and Reprocessing* – EMDR) peut être efficace pour lutter contre les cauchemars que le patient éprouve à la suite d'un traumatisme (Argolo et al. 2015). Ce cinquième exemple illustre l'importance de traiter médicalement les états figés de symptômes préjudiciables. En l'absence d'une telle intervention, la rigidité du réseau resterait importante, empêchant le patient de s'adapter au contexte.

¹¹⁴ Encore une fois, nous insistons sur le fait que l'absence de pathologisation ne signifie pas l'absence de possibilité de prise en charge par le clinicien, qui s'occupe institutionnellement d'un ensemble de conditions dépassant largement les enjeux de la définition d'une espèce psychiatrique telle que nous l'avons proposée. Une telle définition constitue cependant une première étape vers la compréhension de ce dont traite le clinicien dans sa pratique en lui permettant de mieux appréhender celle-ci.

Enfin, pour mieux comprendre l'ensemble de la définition, nous considérons un ensemble de symptômes faisant référence au trouble dépressif caractérisé tel qu'il est décrit dans le DSM-5. Dans cet exemple, tous les éléments de la définition sont considérés. Un jeune patient adolescent demande de l'aide pour une perte d'énergie évoluant depuis six mois et ayant entraîné une anhédonie, qui s'est traduite par une perte d'appétit et une perte de poids importante. Ces symptômes ont entraîné une insomnie d'éveil et participé à l'apparition d'idées de mort récurrentes depuis trois semaines. Aucune mesure prise par lui-même ou ses proches n'a amélioré le préjudice qui sous-tend cet ensemble de symptômes. Chaque symptôme est potentiellement corrélé avec un mécanisme sous-jacent et tous s'influencent mutuellement. Ces effets mutuels dépassent le seuil de tolérance pour cet adolescent et nécessitent une force thérapeutique pour être pris en charge.

Cette définition d'une espèce psychiatrique diffère des autres définitions existantes. Il s'agit d'une approche certes hybride, mais qui s'éloigne cependant de l'HDA de Wakefield et de la définition proposée par Jonathan Tsou (Tsou 2021b, 2022) pour différentes raisons. En effet, premièrement, nous avons proposé une conception fondée sur les symptômes préjudiciables qui ne présuppose pas de variable latente. Deuxièmement, le point de départ de notre définition reposait sur l'abandon des variables latentes (Haslam 2002 ; Kendler et al. 2011 ; Tabb 2019 ; Zachar 2015), en mettant de côté tout construit (et avec cette notion, toute validité de construit) que représentent la catégorie ou la dimension. Troisièmement, nous n'avons cherché qu'à fournir une définition potentiellement utile pour les fonctions cliniques et non pas pour les fonctions non-cliniques. Quatrièmement, nous avons proposé d'ancrer les symptômes préjudiciables à des éléments mécanistiques activés – plutôt que de proposer de discuter la notion de dysfonction. Cinquièmement, la description d'un noyau central nous a permis de regrouper des notions philosophiques, scientifiques et psychométriques au sein d'une même approche. Sixièmement, contrairement à l'approche naturaliste de Tsou, qui soutient l'existence d'espèces biologiques (et cherche

en conséquence à proposer un fondement théorique biologique aux catégories du DSM), nous nous sommes intéressés à un ensemble de symptômes en interaction.

Conclusion

Dans ce chapitre, nous avons discuté des fonctions cliniques de la psychiatrie (diagnostiques, pronostiques et thérapeutiques). Nous avons soutenu que celles-ci nécessitent l'adoption d'une définition de la psychiatrie qui ne se fonde pas sur la notion de variable latente et qui donc doit être distinguée de toute approche par la catégorie / dimension. Nous avons montré que l'approche par le symptôme apparaît adaptée pour répondre à cette nécessité.

Cette approche par le symptôme nous a conduit à discuter une définition théorico-clinique de ce qu'est une espèce psychiatrique utile au clinicien (Cooper 2005 ; De Vreese 2017 ; Reznek 1987). Une telle définition est fondée sur l'existence d'un réseau de symptômes préjudiciables, chacun étant sous-tendu par un mécanisme activé et dont les relations sont figées au sein du réseau. Pour définir une telle espèce psychiatrique utile à la pratique clinique, nous avons identifié deux présupposés relatifs à la dysfonction et au préjudice. Nous avons soutenu que l'identification d'un noyau central permettrait de cibler les symptômes les plus pertinents cliniquement et nous avons précisé les critères (mécanistiques et liés aux préjudices) que devraient comporter les symptômes d'un tel noyau.

Une telle définition théorico-clinique d'une espèce psychiatrique semble pertinente et intuitive au clinicien. Nous avons appelé « intuitive » une définition qui puisse être manipulée par le clinicien sans qu'il n'ait à postuler de variable latente, c'est-à-dire d'entité non-observable dans sa pratique clinique¹¹⁵.

¹¹⁵ Une telle définition ne nécessite pas à proprement parler d'introduire une nouvelle classification au sens d'un « manuel » listant des troubles. Rappelons cependant que, dans ce travail, nous avons utilisé le terme de « classification » comme tout « système d'organisation de composants élémentaires de la psychiatrie ». Dans ce sens, les différentes formes d'une espèce psychiatrique fondée sur une approche par le symptômes (ou « espèce complexe »), variant en fonction des patients, pourraient être considérées comme un tel système d'organisation.

Nous espérons en effet qu'une telle proposition puisse souligner les variations et les contrastes entre individus, en mettant en évidence la diversité au lieu de la réduire à une variable latente, non-observable. Une telle approche se structure au travers des infimes variations entre patients concernant le contenu du noyau central de symptômes et de ses relations¹¹⁶. De la même manière qu'au sein du système solaire, les planètes (en tant que variations d'une même espèce) s'influencent mutuellement par des jeux de gravitation, les symptômes sont les variantes d'un noyau central constituant l'espèce psychiatrique, permettant au clinicien d'enrichir sa compréhension du patient, par glissement d'un symptôme à l'autre, en recueillant ces symptômes. En reconnaissant les « forces de gravitation » les plus importantes, le clinicien peut ainsi savoir sur quel nœud du réseau appuyer pour faire basculer le patient vers le rétablissement. En recherchant une légère variation symptomatique, comme on opère un glissement sémantique entre deux index lexicaux, le clinicien peut actionner une modification du réseau par des changements locaux progressifs, agissant sur les symptômes les uns après les autres. À l'inverse, en l'absence d'adoption par le clinicien d'une telle approche par le symptôme, une forme d'« impatience » conceptuelle a conduit à la volonté de passer d'une classification à l'autre sans étudier les changements locaux de ses composants, un à un, par glissements sémiologiques au cœur d'un noyau central.

Nous avons discuté cette définition théorico-clinique en tant qu'alternative à une définition générique d'une espèce psychiatrique, fondée sur la notion de variable latente, que nous avons étudiée dans les trois premiers chapitres. Une définition fondée sur la notion de variable latente, s'inscrivant dans une approche par la catégorie / dimension, sert de *hub* épistémique pour différentes fonctions non-cliniques. Par exemple, la définition fondée sur la HDA (sous-tendue par la notion de variable latente), intégrée pour une part au DSM-5, pourrait faire office d'un tel *hub* épistémique pour l'ensemble des fonctions non-cliniques que requiert la psychiatrie (par exemple, répondre à des enjeux administratifs ou politiques).

¹¹⁶ On pourrait penser ici à une forme de structuralisme tel que décrit par Lévi-Strauss, qui décrit un système comme un grille de minimales variations glissantes les unes par rapport aux autres, par opposition à un structuralisme – justement « fonctionnaliste » à la Durkheim.

Bien qu'utile en pratique clinique, la définition théorico-clinique ne permet pas de répondre à toutes les exigences non-cliniques décrites dans la première partie de la thèse.

L'existence de deux définitions de ce qu'est une espèce psychiatrique, dans le champ de la psychiatrie, s'inscrit dans un pluralisme définitionnel, non hiérarchique ou horizontal (contrairement à la dualité de Boorse (1977) où la conception pratique est subordonnée à la conception théorique). Ce pluralisme définitionnel serait dit bipartite, c'est-à-dire qu'il tolérerait l'existence d'au moins deux définitions pour la psychiatrie, qui ne devraient pas être réduites l'une à l'autre. En termes philosophiques, un tel pluralisme définitionnel bipartite permet à la psychiatrie de présenter deux types de définitions. Chaque définition est utilisée dans un contexte spécifique¹¹⁷. La notion de préjudice relie ces deux définitions, fournissant un même niveau de seuil pour une espèce psychiatrique, quelle que soit la définition utilisée.

Ces différents développements nous ont permis de proposer une définition d'une espèce psychiatrique en tant que réseau de symptômes préjudiciables figé dans une conformation immuable, inflexible et insensible aux changements de l'environnement et nécessitant une intervention thérapeutique pour être modifié. Les troubles psychiatriques correspondent ainsi à des réseaux de symptômes fortement connectés dont les préjudices cumulés atteignent un seuil de significativité tout en étant sous-tendus par des réseaux de mécanismes activés. Nous espérons que de telles propositions permettront d'avancer vers une meilleure compréhension de la nosologie, des espèces et de la sémiologie psychiatriques.

¹¹⁷ À noter qu'une distinction ne devrait pas conduire à opposer une conception scientifique de la psychiatrie, fondée sur la médecine factuelle, à une psychiatrie axée sur la clinique, fondée sur la subjectivité du clinicien.

CONCLUSION GÉNÉRALE

DES CATÉGORIES AUX SYMPTÔMES

Nous avons tenté d'apporter une modeste contribution au champ éminemment complexe de la philosophie de la psychiatrie en discutant l'intrication de trois principales notions : les classifications, les espèces et les symptômes psychiatriques. Nous avons fait le constat d'une tension, au sein des conceptions et classifications de la psychiatrie, entre leur volonté d'être cliniquement pertinentes et celle d'être pertinentes pour répondre à des fonctions non-cliniques. La diversité des conceptions et classifications en psychiatrie aurait pu découler d'une tentative visant à une concordance entre les aspects cliniques et non cliniques de la psychiatrie.

Nous avons limité le champ des fonctions cliniques au recueil et à l'organisation des symptômes du patient par le clinicien, qui cherche à répondre à des questions diagnostiques et thérapeutiques. Au travers de l'analyse des différentes classifications, nous avons admis que la psychiatrie devait faire face à de multiples fonctions non-cliniques, qui renvoient à tous les enjeux sociologiques, administratifs, financiers, politiques ou communautaires qui entourent les différents acteurs de la psychiatrie.

Dans une première partie de cette thèse, nous avons étudié trois classifications psychiatriques, le DSM, le projet RDoC et la HiTOP. Nous avons tenté de montrer que chacune cherchait à réduire la tension entre ces deux types de fonctions, cliniques et non-cliniques. Nous avons alors mis en évidence que ces conceptions et classifications intègrent comme définition d'une espèce psychiatrique la notion de

variable latente¹¹⁸. Ces conceptions et classifications constituent des approches par la catégorie / dimension. Nous avons supposé que de telles conceptions et classifications et leurs espèces psychiatriques respectives se fondant sur la notion de variable latente (alors appelées « espèces diagnostiques »), pourraient être surtout utiles pour les fonctions non-cliniques. La définition d'un trouble psychiatrique fondée sur l'analyse de la dysfonction préjudiciable, sous-tendue par la notion de variable latente (la dysfonction), constitue un exemple typique d'une approche par la catégorie / dimension. Une telle définition pourrait ainsi servir de plaque tournante (*hub*) épistémique pour l'ensemble des fonctions non-cliniques que requiert la psychiatrie.

Nous avons considéré que la notion de variable latente peinait à être utile en pratique clinique. Nous avons proposé que la définition de ce qu'est une espèce psychiatrique utile pour la pratique clinique puisse ne pas s'appuyer sur la notion de variable latente et se distingue alors d'une approche par la catégorie / dimension. Plutôt, une telle définition renverrait à une approche par le symptôme. Nous avons proposé une définition alternative de ce qu'est une espèce psychiatrique, fondée sur les symptômes. Dans le cadre de cette définition, les symptômes, qui rendent compte cliniquement des préjudices vécus par les patients, constituent les produits de mécanismes activés. L'ensemble organisé de ces symptômes en interactions mutuelles permet de fournir à la pratique clinique une espèce psychiatrique cliniquement pertinente.

Nous avons alors proposé la définition suivante d'une espèce psychiatrique : « Un ensemble de symptômes appartenant au noyau central prototypique de la psychiatrie, organisés en réseau, impliquant au moins un préjudice et dont l'ensemble des préjudices dépasse le seuil de signification socialement

¹¹⁸ Pour rappel, afin de rendre cette conclusion compréhensible indépendamment de ce qui précède, nous avons proposé comme définition de la variable latente toute variable qui ne peut être déduite qu'indirectement à partir d'autres variables observables ou mesurables, qui sert à regrouper plusieurs symptômes en une seule entité, comme c'est le cas d'une catégorie diagnostique ou d'un construit statistique. Pour une part, l'existence d'une variable latente renvoie à ce qui a été désigné, de manière relativement floue, comme une caractéristique essentielle du *modèle médical*. Nous avons cependant évité d'utiliser cette notion afin de mieux approfondir l'analyse épistémologique des classifications, des espèces psychiatriques et des symptômes.

accepté selon les valeurs sociales admises dans une société pour une condition donnée, chacun de ces symptômes étant lié à un mécanisme activé. Un tel réseau de symptômes est figé dans un état qui ne permet pas la diminution spontanée du préjudice relié à ces symptômes sans intervention thérapeutique ».

Une telle définition cliniquement pertinente, fondée sur les symptômes, offre des avantages pratiques pour les cliniciens en permettant des diagnostics, des prédictions, des pronostics et des traitements plus rapides et pertinents. Nous avons montré comment une telle définition améliore la capacité des cliniciens à rapidement identifier d'autres symptômes à partir d'un symptôme initial (en favorisant les « heuristiques » cliniques), notamment en se fondant sur des symptômes qui servent de liens entre différents aspects d'une espèce psychiatrique. Cette approche évite de définir des limites rigides entre les catégories de troubles et limite la surmédicalisation en utilisant un seuil de préjudice. De plus, elle s'aligne avec la médecine de précision en se concentrant sur les mécanismes physiopathologiques des symptômes plutôt que sur les troubles dans leur ensemble. Enfin, elle cherche à raffiner la validité de contenu en structurant les symptômes au sein de réseaux, ouvrant ainsi de nouvelles perspectives en philosophie de la psychiatrie.

Assumer l'adoption d'au moins deux définitions pour la psychiatrie semble nécessaire pour répondre aux enjeux de la pratique clinique et aux enjeux non-cliniques de la psychiatrie. Un tel constat ouvre la voie à une forme de réconciliation entre différentes conceptions de la psychiatrie, accueillant une forme de tolérance concernant la multiplicité de définitions de ce qu'est un trouble psychiatrique. À l'avenir, rien n'empêcherait que d'autres définitions de ce qu'est une espèce psychiatrique soient également distinguées. Par exemple, proposer une définition d'une espèce psychiatrique utile pour les patients, distincte d'une définition utile pour les politiques et les décideurs publics, pourrait être nécessaire. Un tel travail soutient ainsi les promesses de dialogue entre philosophie, nosologie, sociologie, histoire et psychiatrie clinique et, nous l'espérons, favorisera les discussions en philosophie de la psychiatrie composée elle-même de multiples communautés.

ANNEXES

Annexe 1 – Dépendance au chemin et régime des promesses dans la nosologie psychiatrique

Rationnel de cette annexe : Il s'agirait de comprendre en quoi la dépendance au chemin et le régime des promesses permettraient de mieux comprendre la dynamique des classifications psychiatriques.

Développant une approche par les systèmes d'innovation dans un tout autre champ, Vanloqueren et Baret (2009) ont identifié un certain nombre de déterminants de l'innovation (c'est-à-dire les facteurs qui influencent les choix de recherche) dans le cadre (des priorités de financement) des systèmes de recherche (qui intègrent de manière symétrique l'objet de science comme point de passage obligé et les acteurs du réseau) (Vanloqueren & Baret, 2009). De manière similaire, dans cette annexe, nous nous appuyerons sur cette notion de dépendance au chemin pour expliquer le phénomène d'itération épistémique (voir Liebowitz et Margolis, 2000¹¹⁹) (Liebowitz & Margolis, 2014). Nous soutenons dans cette annexe que la dépendance au chemin peut permettre de mieux comprendre l'influence d'une promesse sur le renforcement de la légitimité et de la crédibilité d'une classification psychiatrique (Pierson, 2000; Vanloqueren & Baret, 2009). En effet, la dynamique d'auto-renforcement des promesses et des classifications par l'intermédiaire de la légitimité et de la crédibilité pourrait être partiellement éclairée par la notion de dépendance au chemin. Celle-ci renvoie au fait que les voies de recherche alternatives, c'est-à-dire celles qui ne sont pas entretenues à un stade précoce par les acteurs du réseau, deviennent des « options mortes », pour reprendre le terme de William James dans *The Will to Believe* (1896)¹²⁰ (James, 1897). Certaines promesses (et donc classifications) sont bloquées dans leur développement en raison de l'irréversibilité des trajectoires économiques et techniques d'autres promesses. Par exemple, sur un plan économique, l'appel de fonds majeur pour le séquençage du génome humain dans les années 1990 aurait pu limiter le développement d'autres heuristiques potentielles comme la recherche sur les expositions environnementales des maladies. Inversement, le développement de certaines promesses (ou classifications) a fait progresser la science dans une voie (un chemin) qui ont incité les chercheurs, laboratoires et organismes à poursuivre dans cette voie plutôt que d'en recommencer une autre, par exemple dans la recherche des facteurs environnementaux (Rappaport, 2011; Wild, 2005). Il y a donc une dépendance de la science (et des classifications) au chemin initié par une promesse qui a conquis la légitimité (et la crédibilité) de la communauté. Le cycle de la promesse auto-entretenant la classification s'inscrit dans ce chemin déjà parcouru¹²¹ (Ruphy, 2017). En ce qui concerne les classifications, à l'avenir, il est plus probable que la recherche en neurosciences utilise le cadre du projet RDoC, qui a développé sa légitimité sur un réseau d'acteurs publics et scientifiques (et possède des fonds propres), par rapport à tout nouveau projet taxonomique neuroscientifique intégratif – du moins tant que l'éventuelle validité supplémentaire de ce nouveau projet ne conquerrait pas la légitimité des acteurs déjà impliqués dans le réseau. De même, le développement nosologique de la psychanalyse se baserait plutôt sur le *Psychodynamic Diagnostic Manual* (PDM) que sur une nouvelle classification et une classification statistique se fonderait plutôt sur les progrès inaugurés avec la HiTOP. En d'autres termes, il existe une certaine dépendance aux conditions initiales des classifications en cours de développement.

¹¹⁹ Les auteurs écrivent : « La dépendance au chemin peut être faible (l'efficacité du chemin choisi est liée à certaines alternatives), semi-forte (le chemin choisi n'est pas le meilleur mais ne vaut pas la peine d'être corrigé), ou forte (le chemin choisi est très inefficace, mais nous ne pouvons que difficilement le corriger) ».

¹²⁰ Précisément, James distinguait des « *live options* » possédant une résonance avec l'expérience et perçues comme de réelles possibilités et des « *dead options* », n'ayant que peu de résonance imaginative et n'étant ainsi pas développées et soutenues.

¹²¹ En utilisant un autre vocabulaire, il faut une énergie supplémentaire pour donner de la légitimité et de la crédibilité à une classification qui n'est pas le chemin « dominant », par rapport à l'énergie dépensée pour faire progresser la classification déjà légitimée et crédibilisée par la communauté clinique et scientifique.

L'application de la dépendance au chemin au régime des promesses amène au moins trois implications. Premièrement, elle renforce l'argument selon lequel les invariants traditionnellement cités pour expliquer le succès d'une classification (par exemple, la validité, la fiabilité, l'applicabilité ou l'efficacité clinique) sont secondaires. Dans ce sens, il apparaît autant souhaitable d'isoler les invariants ayant permis l'extension de la légitimité, de la crédibilité et du déploiement d'une classification que d'identifier des facteurs contingents liés à la tradition, à la rhétorique de rationalité, au charisme des acteurs, à l'intégration cohérente dans une communauté scientifique, aux outils développés ou aux collaborations scientifiques et cliniques. Le développement de la nosologie psychiatrique aurait potentiellement pu être plus fructueux si le chemin emprunté avait été un autre – voies tangentielles que la science ou la clinique n'ont pu emprunter en raison d'une dépendance (même partielle) au chemin parcouru¹²².

De plus et pour nuancer cette dernière affirmation, cette irréversibilité n'est pas totalement déterministe : comme l'ont développé Garud et Karnoe dans leur ouvrage *Path dependence and creation* (Garud et al., 2010), la création de nouveaux chemins aide à identifier les ressources nécessaires pour créer une nouvelle dépendance au chemin (voir également Hasok Chang, 2004, pour un inventaire des avantages liés à l'exploration dynamique de voies parallèles – dans la lignée de Feyerabend dans *Against Method* (1975)) (Chang, 2004; Feyerabend, 1975).

Enfin, pour rebondir sur le point précédent, la possibilité de création de nouveaux chemins pourrait être nuancée. En effet, ceux-ci pourraient conduire à un appauvrissement épistémique. La réduction des futurs peut être nécessaire : un nombre trop important de promesses pourrait conduire à une dilution de l'unité de la discipline, des fonds et des ressources humaines, se répercutant sur la vision des politiques et du grand public. À l'heure actuelle, le pluralisme nosologique (intégré au cadre plus global de la multiplicité des perspectives) entraîne la psychiatrie vers cette voie de la désunicité, qui peut certes lui être nuisible pour certaines de ses priorités (comme le fait de convaincre les politiques publiques de la scientificité de sa pratique), même si elle constitue l'une de ses principales forces quand il s'agit de comprendre le patient en pratique clinique ou un phénomène neuroscientifique.

Nous avons donc vu dans cette annexe comment la dépendance au chemin des classifications limitait la contingence des facteurs sociologiques, constituant une orientation régulatrice de la dynamique scientifique.

¹²² En étant plus radical et en donnant encore plus de poids à la dépendance nosologique au chemin, on pourrait dire que celle-ci fait passer d'un potentiel pluralisme nosologique, qui serait un pluralisme de promesses, à une exigence d'utilisation de la classification déjà en place (pensons au DSM). Nous ne développerons cependant pas cette argument qui, selon nous, donne trop de poids à l'inertie d'une découverte scientifique pour négliger la pression de l'innovation (même si elle est implicite).

Annexe 2 – Approche réseau et clusters de propriétés homéostatiques

Rationnel de cette annexe : Il s’agirait de comprendre en quoi la définition d’une espèce psychiatrique fondée sur la notion de réseau s’intègre et se distingue de la conception des clusters de propriétés homéostatiques.

Dans l’approche réseau, les espèces psychiatriques ne sont pas unifiées par une essence sous-jacente partagée (la variable latente, ou cause commune), mais par des relations causales entre des symptômes (ou d’autres variables non sémiologiques). De fait, les réseaux de symptômes peuvent être compris comme des *clusters de propriétés homéostatiques sémiologiques*, unifiés par des relations causales (Boyd, 1988, 1991, 1999; Hacking, 1991) (ou comme des « clusters de propriétés mécanistiques sémiologiques », en reprenant les termes de (Kendler et al., 2011)). Sous ce modèle, les troubles psychiatriques sont définis comme des ensembles de symptômes suffisamment unifiés par leurs relations causales pour soutenir l’induction, l’explication et la prédiction (Hofmann et al., 2016; Whewell, 1860). Le réseau constitué de symptômes constitue lui-même un mécanisme *stipulatif et nominal* (non physiologique) composé de symptômes en interaction. Gupta (2019) différencie en effet les définitions stipulatives (qui stipulent les significations des termes, sans tenir compte des significations antérieures ou établies des termes), descriptives (qui précisent les significations des termes en fonction des usages établis) et explicatives (qui spécifient les significations des termes en respectant certaines des utilisations les plus centrales d’un terme) (Gupta, 2019).

Quelle est vraiment la conception de Boyd et comment s’applique-t-elle à la psychiatrie ? Tabb (Tabb, 2019) (p. 2185) affirmait que la proposition de Boyd ne serait qu’un « récit métaphysiquement permissif », qui ne nécessite aucun « mécanisme sous-jacent métaphysiquement exigeant »¹²³. Que proposait Boyd, à l’origine ? Dans un article de 1999, il écrivait qu’une famille de propriétés cooccurrentes « est le résultat de ce que l’on peut qualifier d’homéostasie : soit la présence de certaines propriétés tend à favoriser la présence d’autres [i], soit il existe des mécanismes sous-jacents qui tendent à maintenir les propriétés de cette famille [ii], ou les deux [iii] » (Boyd, 1999) (p. 143). Nous acceptons donc que la présence de certains symptômes tend à favoriser la présence d’autres symptômes (i) et pouvons admettre qu’il existe des mécanismes (ii), bien que nous restions agnostiques à ce sujet d’un point de vue nosologique.

Boyd, en soutenant sa conception des espèces en tant que clusters de propriétés homéostatiques, défend une théorie non essentialiste des espèces naturelles (Boyd, 1991, 1999). L’intérêt d’une telle espèce tient notamment à sa capacité à produire des inférences projetables (dites « ampliatives »). Les principales caractéristiques des clusters de propriétés homéostatiques sont les suivantes (Boyd, 1999) (p. 143) : i) il existe une famille de propriétés qui sont groupées de manière contingente ; ii) leur cooccurrence est le résultat de l’homéostasie, c’est-à-dire que la présence de certaines propriétés tend à favoriser la présence d’autres propriétés. Boyd ajoute la notion de « terme générique » [*kind term*], qui correspond à l’ensemble des éléments du regroupement homéostatique et définissant ce regroupement comme une espèce – permettant d’expliquer la réussite des inférences projetables. Ainsi, pour la psychiatrie, les clusters de propriétés homéostatiques permettent de mieux comprendre un trouble comme la dépression (Borsboom &

¹²³ Dans la littérature, il existe un désaccord quant à savoir si l’une des propriétés ou des mécanismes sous-jacents des clusters de propriétés homéostatiques doit être intrinsèque. Les philosophes de la biologie soulignent l’importance des propriétés et des mécanismes relationnels (par exemple, les relations phylogénétiques) maintenant la stabilité des groupes de propriétés associés à une espèce (Boyd, 1999; Ereshefsky, 2009; Wilson et al., 2007), d’autres (par exemple, (Griffiths et al., 1999; Millikan, 1999)) insistent sur le fait que certaines propriétés intrinsèques (par exemple, les propriétés génétiques) sont nécessaires.

Cramer, 2013; Kendler et al., 2010). Celle-ci est particulièrement hétérogène¹²⁴ et il est difficile de tracer une frontière discrète entre normal et pathologique – expliquant les taux élevés de comorbidité de la dépression avec d’autres troubles comme le trouble anxieux. Du point de vue des clusters de propriétés homéostatiques, il faut s’attendre à ce que les individus d’un groupe de propriétés A (par exemple, les symptômes de la dépression du DSM) se retrouvent souvent dans un autre groupe B (par exemple, ceux du trouble anxieux généralisé), car les groupes partagent des propriétés déterminantes : ils englobent tous deux les « problèmes de sommeil », la « fatigue », les « problèmes de concentration » et « l’agitation psychomotrice ». Les syndromes se chevauchent substantiellement et les personnes présentant un certain ensemble de symptômes seront souvent décrites de la même manière par des diagnostics différents. Cela explique également les problèmes de manque d’efficacité des études sur les traitements médicamenteux, le manque de marqueurs biologiques de la dépression et la faible fiabilité et validité des diagnostics : des seuils arbitraires catégoriels ne reconnaissent pas correctement l’hétérogénéité de tels troubles, ni les relations entre les symptômes – ce que permet la conception d’une espèce complexe comme décrite au sein de l’approche réseau, fondée sur les clusters de propriétés homéostatiques.

Compte tenu de cette conception des espèces psychiatriques comme espèces complexes, la théorie des réseaux de symptômes peut fournir une base pour le développement des classifications psychiatriques. Contrairement aux propositions récentes basées sur des données neuroscientifiques (comme les RDoC), qui décrivent des entités élémentaires renvoyant à des espèces naturelles (Cuthbert, 2014), ou aux propositions fondées sur des bases psychométriques et statistiques (comme la HiTOP) (Krueger et al., 2018), les réseaux de symptômes sont à la fois disruptifs et conservateurs : ils conçoivent une espèce constituée d’éléments en interaction, tout en conservant partiellement les unités élémentaires du DSM (les symptômes – voir principe 2 de la théorie des réseaux) – du moins comme point de départ (Haslam, 2002; Tabb, 2019).

¹²⁴ Ce constat est vrai au niveau des symptômes, mais également de la biologie. Une étude portant sur plus de 12 000 personnes a révélé que les personnes atteintes de schizophrénie présentaient un taux inhabituellement élevé de mutations ultra-rares – et que celles-ci étaient souvent uniques à une seule personne ((Brainstorm Consortium et al., 2018), *Science*).

Annexe 3 – Définition du noyau central

Rationnel de cette annexe : Il s'agirait de mieux comprendre les implications de la relation entre le « noyau central » et la « périphérie » d'une espèce psychiatrique.

Le noyau central peut être considéré comme un module (i.e. un sous-ensemble de symptômes) constitué des symptômes les plus influents sur le reste du réseau (renvoyant à la notion de centralité). Par respect pour le principe 2 de la théorie des réseaux de symptômes, pour des raisons de stabilité nosologique et d'organisation des soins, le contenu du noyau central devrait être constitué des symptômes décrits dans les catégories traditionnelles des classifications. En effet, une classification comme le DSM range les symptômes au sein de critères formels sémiologiques, monothétiques ou polythétiques (comme le « critère A. »). Le noyau central correspond donc aux symptômes les plus influents sur les autres symptômes du réseau et il est constitué de ces symptômes traditionnellement décrits.

Autour de ce noyau central, nous avançons l'existence d'un ensemble de symptômes périphériques. Ce réseau périphérique permet de donner des indications sur l'espèce psychiatrique sans pour autant que ses composants soient nécessaires ou suffisants dans la définition de ce qu'est un trouble. Il est donc constitué d'éléments possédant une ou plusieurs de ces trois propriétés : i) il s'agit soit de symptômes moins « importants » en termes d'influence sur le réseau, mais appartenant au syndrome clinique tel que défini selon le répertoire psychiatrique (par exemple, les troubles de l'appétit dans l'épisode dépressif peuvent être analysés comme moins centraux que l'anhédonie, en se fondant sur un réseau d'arguments de différentes natures comme les données probantes, l'histoire de la psychiatrie, la modélisation, etc.) ; ii) soit des caractéristiques du trouble qui sont nécessaires à la critériologie, mais qui ne doivent pas être confondues avec les indicateurs d'un trouble clinique, comme par exemple le « critère de significativité clinique » (Cooper, 2013; Spitzer, 1999b) ; iii) soit d'autres spécificateurs, tel que l'âge, le sexe, ou d'autres données démographique, mais également d'autres variables plus instables sur le plan de la preuve scientifiques, tels que des biomarqueurs. En effet, concernant ce dernier point (iii) et comme nous l'avons soutenu précédemment, des éléments et variables environnementaux, démographiques ou biologiques peuvent être intégrés au réseau.

1. Implications pour le clinicien et la nosologie

La définition d'un noyau central spécifique à chaque trouble psychiatrique est d'une grande importance, car c'est l'évaluation de ces symptômes les plus centraux par le clinicien qui conditionne pour une grande part la prise en charge du patient. Un tel noyau central peut suffire au clinicien à poser le diagnostic, tout en répondant aux différents impératifs des instances et institutions en termes administratifs, politiques, législatifs, économiques ou sociologiques. En effet, l'identification d'un noyau central, qui serait distingué d'un réseau périphérique, soutient le principe de conservatisme prudentiel qui n'empêche pas une classification (Demazeux, 2015) de répondre à de nombreux et divers enjeux (Sullivan, 2017). En résumé, un tel noyau central pourrait aider à relever les défis de : i) servir de structure catégorielle stable pour les institutions cliniques, la recherche, les administrations, les politiques et les assurances; ii) être modifié uniquement lorsqu'une modification de fond de la classification est envisagée ; iii) constituer le « centre de gravité » d'un trouble en vue d'être validé par son caractère prédictif (possibilité de fournir un traitement) et de sa structure pronostique (définition du trouble selon son évolution) (Fusar-Poli et al., 2018).

En mettant de côté les éléments non centraux ou non cliniques (comme le « critère de significativité clinique » ou les spécificateurs d'âge, de fréquence, etc.), le réseau périphérique permet de différencier les critères diagnostiques et les éléments indicateurs d'un trouble pathologique (Bolton, 2013b). De plus, un tel réseau périphérique peut être adapté plus soupagement lors des révisions nosologiques, enjeu compatible avec la volonté d'un « DSM-live » (American Psychiatric Association, 2013), mis à jour de manière dynamique au fil du temps. Ensuite, à l'image de ce qui a été réalisé dans des études comme CONVERGE (Cai et al.,

2015), un réseau de ce type permet d'étudier en recherche des symptômes n'appartenant pas au DSM (Kendler et al., 2018) et plus généralement au « répertoire psychiatrique » traditionnel. Par exemple, il peut permettre d'intégrer des *Core Outcome Sets* (COS), symptômes jugés les plus pertinents par les patients pour définir leur propre condition de santé, ou d'intégrer des *Patient Related Outcome Measures* (PROM), qui recueillent par des méthodes numériques les domaines clés permettant de mesurer la santé relativement aux points de vue des patients (Black, 2013). Le réseau périphérique peut permettre d'accueillir ces « nouveaux symptômes », constituant une étape transitionnelle avant leur intégration potentielle dans le noyau central si leur robustesse et leur validité venait à être confirmée au fil du temps.

2. Quelle sorte d'espèce psychiatrique constitue ce noyau central ?

En reprenant l'image du réseau comme un « répétable » selon Cooper (2012), si nous prenons différents cas de troubles psychiatriques et les modélisons sur un espace multidimensionnel, par exemple avec une analyse en clusters, nous trouverons un ou plusieurs clusters similaires et plus importants que d'autres, ces clusters pouvant être considérés sur le plan conceptuel comme le noyau central du réseau. Une telle détermination n'empêche d'ailleurs pas la validation externe des clusters (ou « communautés » ou « modules », dans le langage des réseaux) sur la base de la réponse au traitement, des antécédents familiaux, des corrélats démographiques, ou tout autre valideur (pour une discussion récente sur la hiérarchisation des validateurs en psychiatrie, voir Solomon et Kendler, 2021, *The Problem of Aggregating Validators for Psychiatric Disorders* – (Solomon & Kendler, 2021)).

Le noyau central est également en accord avec la revendication de Tabb (2019) plaidant pour une meilleure « discrimination diagnostique », selon laquelle « les philosophes de la psychiatrie ne devraient pas se concentrer exclusivement sur les espèces diagnostiques », c'est-à-dire les catégories traditionnelles de troubles mentaux (comme l'ont fait notamment Murphy, 2006 ou Cooper, 2013). Cependant, la notion de noyau central évite de tomber dans l'écueil de Tabb (2019), dont le risque est celui du relativisme sur les espèces psychiatriques : un tel noyau fournit un fondement solide et stable pour discuter des troubles psychiatriques, en acceptant le pluralisme nosologique.

Comme nous l'avons vu avec la présentation de l'approche réseau, le noyau central peut être également considéré comme une espèce complexe (Boyd, 1988) : il est constitué d'éléments hétérogènes en interaction, potentiellement évolutifs, au sens du cluster de propriétés homéostatiques.

De plus, en reprenant la définition de Zachar (2000), il est une espèce pratique, comme nous le détaillerons mieux par la suite : il est constitué de certains symptômes (et pas d'autres) *du fait même* de leur manipulabilité diagnostique, pronostique, prédictive ou thérapeutique potentielle offerte au clinicien ou chercheur. Notons que ce n'est pas parce que le clinicien les manipule qu'ils sont centraux, mais bien du fait que ces symptômes sont centraux, le clinicien et le chercheur s'y intéressent, portent dessus leur attention et raffinent leurs outils pour mieux les appréhender et modifier. Nous en viendrons à critiquer la distinction de Zachar (2000, 2014) entre les espèces naturelles et les espèces pratiques : les clusters de propriétés homéostatiques offrent une description naturelle, capable de fournir un modèle stable de symptômes (et autres variables) en interactions, tout en permettant de fournir un certain « succès pragmatique » tel que l'a détaillé Zachar.

Enfin, le noyau central peut également être considéré comme une espèce floue (Haslam, 2002), puisque on retrouve une graduation de centralité : certains symptômes sont plus centraux que d'autres, puisqu'on peut pondérer de différentes manières les mesures de centralité.

Pour aller plus loin, certains marqueurs objectifs (par exemple, biologiques, génétiques, etc.) peuvent être placés dans le réseau dynamique périphérique dès lors que des données probantes à leur sujet pourraient être fournies. Cette intégration périphérique pourrait s'appuyer sur la matrice des RDoC, qui fournit des marqueurs neuroscientifiques validés par des données scientifiques en constante évolution en fonction de nouvelles données scientifiques (Cuthbert, 2014). Intégrer des marqueurs de ce type sans pour autant les considérer comme des éléments centraux du trouble peut être d'une grande utilité pour la recherche (voire,

à terme, la clinique psychiatrique) tout en offrant une grande prudence concernant les risques de refonte de ce qui constitue le cœur de la psychiatrie.

3. Implications du noyau central du réseau pour la stabilité de la nosologie

Dans cette définition fondée sur l'approche en réseau, la prise en compte du contexte et des intérêts épistémiques pourrait conduire à tomber dans l'écueil du relativisme et, en pratique, à une instabilité des éléments du réseau. Cependant, les symptômes constituant le noyau d'un réseau restent stables. En effet, ces symptômes qui constituent le noyau du réseau sont majoritairement des symptômes issus du répertoire traditionnel de la psychiatrie (et donc principalement du DSM). Comme nous l'avons vu, ce sont les symptômes périphériques qui peuvent évoluer en fonction du contexte, de la question du chercheur, du patient, mais aussi, à plus long terme, des progrès scientifiques. La périphérie peut évoluer et s'adapter de manière dynamique en fonction des progrès de la psychiatrie de précision, des théories et questions posées par le clinicien ou chercheur et du contexte environnemental. Cette idée d'un noyau central et d'une périphérie mouvante recoupe partiellement la notion de programme de Lakatos (1970), qui est constitué d'une théorie principale (le noyau dur central) et de théories auxiliaires (la ceinture de protection périphérique) (Lakatos, 1980). Comme nous l'avons dit dans la thèse, cette notion de centralité renvoie à de nombreuses conceptions en psychiatrie, de la notion de symptôme de premier rang dans la psychose (Schneider, 1959) à celle de « répétables » chez Cooper (2012).

4. Noyau central en philosophie de la médecine

Nous pouvons discuter au moins trois implications de cette notion de noyau central pour la conception des troubles et la nosologie, en nous fondant sur les développements de la philosophie de la médecine.

Itération épistémique. Premièrement, notons qu'une telle conception est en accord avec le modèle d'itération épistémique du DSM, selon lequel des révisions successives (en périphérie) permettent d'atteindre une complexification croissante des classifications (Chang, 2004 ; Kendler, 2009). La présence d'un système périphérique garantit : i) un réalisme minimal (la cible visée ne se modifie pas trop au fil des itérations épistémiques), ii) une stabilité de l'itération (il n'y a pas de rupture paradigmatique lors du processus) tout en maintenant iii) et un consensus minimal sur les conditions initiales.

Pragmatisme. Deuxièmement, on retrouve dans cette différenciation entre centralité et périphérie une forte implication pragmatique : le noyau central est avant tout défini de manière pragmatique (comme les catégories du DSM) ; la périphérie est moins pragmatique, mais sert mieux la validité de construit.

Conceptions idiographiques et nomothétiques. Troisièmement, la notion de noyau central renvoie à une conception nomothétique. Le noyau central est en effet commun à un ensemble de patients, puisque ses variables (symptômes) ont été identifiées au cours des observations historiques de la psychiatrie (intégrées au sein des versions du DSM), puis vérifiées par les analyses en réseau elles-mêmes. Ce noyau central est donc défini sur le plan populationnel ; il est donc dit nomothétique. Qu'en est-il pour un patient donné ? Ce noyau central, dit alors idiographique, pourrait être différent. Cette modélisation aurait d'autres utilités, par exemple, observer l'évolution des variables centrales au fil de l'histoire naturelle de la maladie du patient. Le choix pragmatique des variables du noyau central pour un patient donné dépend des caractéristiques de l'individu, du contexte et de la question (par exemple, thérapeutique) du praticien. Souligner la nécessité d'une architecture commune pour un trouble psychiatrique et accepter qu'il existe une certaine forme d'invariance en termes d'expressivité des troubles (le noyau central), constitue une condition nécessaire pour que la prévention, le diagnostic et la thérapeutique puissent être mis en place. À l'inverse, c'est parce qu'il existe une périphérie mouvante et constituée de données précises et propres à un patient qu'une psychiatrie personnalisée et de précision peut être développée (Gauld et al., 2021). La notion de noyau

central permet donc de justifier la double valence de la conception que nous développerons : à la fois conception générale nomothétique se rapprochant de la pathologie [*disease*] et conception individuelle idiographique se rapprochant de la maladie [*illness*], cette dernière comportant une attention particulière pour les paramètres phénoménologiques, mécanistiques, physiologiques et la souffrance individuelle à l'échelle des symptômes (Engelhardt, 1973, 1974, 1981)¹²⁵.

Pour conclure, nous soutenons donc que la définition d'une catégorie du DSM pourrait inclure : i) des symptômes centraux (noyau central) déterminés sur des grandes quantités de données cliniques et psychométriques et selon des mesures valides issues de théories psychopathologiques robustes, garantissant la fiabilité de la classification ; ii) et des critères potentiellement dynamiques (réseau périphérique) cherchant à améliorer (modiquement) la validité de la discipline psychiatrique grâce à l'intégration de données issues de la médecine fondée sur les données probantes.

¹²⁵ L'analyse se complique foncièrement lorsqu'on essaie de relier ces deux vues, ce que nous ne tenterons pas d'argumenter ici. Elle reviendrait à poser la question : « La maladie existe-t-elle ou n'y a-t-il que des personnes malades », soulevant des difficultés ontologiques que Peirce, entre autres, a discuté en 1965 en renvoyant aux conséquences pratiques qui en découlent : « C'est la somme des conséquences [des cas particuliers – idiographique] qui constitue la signification entière de cette conception [générale – nomothétique] ». Ainsi, par exemple, Kirmayer et Young (1999) écrivaient que « les gens raisonnent sur les troubles psychiatriques sur la base d'analogies et ils s'appuient sur des prototypes et des exemples » (Kirmayer & Young, 1999) (p. 451) ; Wakefield (1999), avec son « essentialisme de la boîte noire », proposait que les « propriétés de surface d'un ensemble » informent le « noyau conceptuel », qui restait indépendant du raisonnement des individus (Il écrivait : « Si je me fie à l'image d'une Pontiac pour raisonner sur les voitures, cela ne veut pas dire que mon concept de base de la voiture fait référence aux Pontiac » (1999).

Annexe 4 – Le spectre d’essentialité

Rationnel de cette annexe : Le spectre d’essentialité pourrait répondre à une certaine forme d’insatisfaction vis-à-vis des réseaux de symptômes décrits dans la seconde partie de ce travail.

Dans cette annexe, en nous fondant sur l’exemple du DSM et du projet RDoC, nous allons introduire la notion de « spectre d’essentialité ». Ce recours au spectre d’essentialité provient d’une relative insatisfaction vis-à-vis des réseaux de symptômes décrits dans le travail de thèse : doit-on vraiment laisser à l’influence causale entre symptômes (préjudiciables) le poids des explications d’un trouble ? Ne peut-on *dans certains cas et pour répondre à certaines questions* envisager l’existence d’une variable latente, plus ou moins essentialisée, qui puisse expliquer ce regroupement de symptômes ? N’y-a-t-il pas certaines conditions pour lesquelles il serait plus fructueux de penser l’existence de regroupements de symptômes en termes d’essence sous-jacente ? Et si c’est le cas, comment modéliser ensemble certains tableaux cliniques mieux expliqués par des relations causales entre symptômes (par exemple, l’anxiété entraînant des troubles du sommeil, qui entraînent une baisse de l’humeur) avec d’autres tableaux cliniques mieux expliqués par des facteurs communs (par exemple, les symptômes du syndrome de Williams, tous liés à une mutation génétique) ? En effet, comme cela est décrit dans un certain type de pluralisme (saupoudré), pour certains troubles, dans certains contextes et pour un certain type de patients, les mécanismes environnementaux contribuent fortement à leur expression (par exemple, il existe des expressions fortement dépendantes de la culture) ; pour d’autres troubles, dans d’autres situations, les mécanismes biologiques déterminent plus fortement leurs caractéristiques causales.

Interroger le spectre d’essentialité permet de mieux comprendre l’influence mutuelle des symptômes : certains troubles à forte essentialité sont mieux décrits comme des réseaux de modifications biologiques (entraînant des préjudices). À l’inverse, d’autres troubles, dits « à faible essentialité », sont mieux décrits comme des réseaux de préjudices, relativement indépendamment des mécanismes. Il faut dans tous les cas garder à l’esprit que le spectre d’essentialité fournit une *probabilité* d’identification du mécanisme (il est moins probable d’identifier les mécanismes biologiques à l’origine d’un deuil que les mécanismes biologiques à l’origine d’hallucinations). Il est important de comprendre que le long de ce spectre d’essentialité, il peut être plus ou moins *espéré* trouver une influence causale entre symptômes. L’intérêt n’est pas l’identification empirique de cette variable commune (essentialiste), mais bien la possibilité qu’une telle variable puisse mieux expliquer (causalement) l’agrégation des symptômes¹²⁶. Dans cette annexe, nous verrons comment la psychiatrie computationnelle s’adresse aux symptômes relativement indépendamment de ces classifications, pour montrer leurs relations avec des mécanismes sous-jacents qui en expliquent la survenue.

En utilisant les outils de la psychiatrie computationnelle, nous en viendrons nécessairement à devoir prendre en compte la dynamique temporelle des troubles et leur contexte. Nous étudierons ainsi l’importance de la phénoménologie (intentionnalité) et de la dynamique (stadification) des troubles psychiatriques. Nous utiliserons l’exemple de la HiTOP pour explorer ce potentiel dimensionnel, ce qui nous conduira à discuter un modèle de conception des troubles psychiatriques dimensionnel et évolutif : le modèle de stadification (ou *staging*). Cette annexe permettra ainsi de comprendre que la psychiatrie computationnelle véhicule un fort potentiel de complexification des systèmes de compréhension et des modèles actuels des troubles psychiatriques. Les catégories du DSM n’ont pu être corrélées à des substrats biologiques et sont apparus trop hétérogènes pour permettre de tirer des conclusions les reliant à de quelconques mécanismes. Le modèle de la parésie générale des aliénés, aujourd’hui la syphilis à un stade avancé, constitue un exemple de troubles

¹²⁶ Rappelons qu’au fil des développements de ce travail de thèse, les classifications psychiatriques sont utilisées comme des vecteurs de discussion, en tant qu’elles véhiculent de manière concrète les concepts, notions, valeurs et styles de raisonnement des systèmes qui les sous-tendent.

psychiatrique de haute essentialité : au début du 20^e siècle, on a découvert que ce trouble était lié à la bactérie *Treponema pallidum* (un spirochète), qui produit une atrophie frontotemporale. Mais comme l'a noté Kendler (2005), « nous ne pouvons plus nous attendre à des découvertes 'de type spirochète' » (Kendler, 2005). Très précocement dans son histoire, l'un des principaux défis de la psychiatrie computationnelle a été de s'extraire de la dépendance des catégories et des notions véhiculées par le DSM. Par exemple, comme nous l'avons vu dans l'annexe précédente, les troubles psychiatriques pourraient être considérés comme des réseaux en interaction dynamique et variant de manière dimensionnelle (Borsboom, 2008, 2017).

Il faut d'abord rappeler les écueils de la pensée essentialiste concernant la psychiatrie : certes, le DSM est sous-tendu par l'idée que ses catégories ne sont que des construits, mais ces construits peuvent être rassemblés sous une définition générale qui pose une forme d'essentialisme (de la boîte noire – (Wakefield, 1999)). Ce qui permet de justifier l'existence d'un tel essentialisme est que la fonction de ces mécanismes supposés est décrite de manière évolutionniste. Autrement dit, c'est parce que ces mécanismes ont une fonction sélectionnée par l'évolution qu'ils rendent compte d'un agrégat de symptômes (formant un trouble psychiatrique). Cette forme d'essentialisme, qui se rapproche du « paramètre fictif » [*essence placeholder*] de Medin et Ortony (Medin & Ortony, 1989), postule qu'un dysfonctionnement sous-jacent, non nécessairement identifié sur le plan empirique, sous-tend les regroupements phénotypiques formant les catégories diagnostiques. Cette pensée essentialiste est fructueuse pour répondre à de nombreux objectifs, pour la plupart non cliniques (voir (Brick et al., 2020) pour une discussion sur le sujet). En reprenant un exemple devenu célèbre, il est possible d'expliquer l'assortiment relativement indépendant des gènes lors de la méiose sans nous préoccuper des mécanismes chimiques sous-jacents exacts (Mitchell, 2009). Nous pouvons prédire la distribution des traits hérités des organismes via la génétique classique sans rien savoir sur l'ADN, ou prédire la réponse qu'un calcul informatique fournira pour un problème arithmétique, tout en restant ignorant sur les propriétés électriques des semi-conducteurs.

Il en est de même pour des caractéristiques superficielles d'un phénomène comme les « comportements », à l'échelle macroscopique. Les explications fournies par l'étude des comportements, au niveau supérieur, permettent des explications et des prévisions convenables pour les cliniciens et (certains) chercheurs. Autrement dit, il serait possible (voire souhaitable) d'adopter une relative indifférence vis-à-vis des explications des niveaux inférieurs. Ce qui est important n'est pas tant de savoir si un phénomène donné de niveau supérieur est réellement mis en œuvre dans le monde de différentes manières physiques (voir notamment (Kim, 1993)), mais de comprendre comment un niveau peut être différemment informatif par rapport à un autre (voir notamment la notion d'émergence¹²⁷). En nous fondant sur la possibilité de concevoir les troubles psychiatriques comme des systèmes complexes, tel que développé notamment dans l'annexe précédente (Kendler et al., 2011), nous avons affirmé qu'une telle essence ne devait pas nécessairement sous-tendre chaque trouble. Il convenait plutôt d'identifier autant de mécanismes que de symptômes.

La psychiatrie computationnelle raffine encore cette proposition : elle permet d'avancer que la *possibilité* d'identification d'une essence sous-jacente à chaque symptôme varie le long d'un « spectre d'essentialité » (Hitchcock et al., 2021). Autrement dit, nous assumons que certains troubles et dans certaines conditions (nous allons détailler cela par la suite) pourraient bénéficier de la détermination de paramètres latents (neurobiologiques) pour soutenir le diagnostic et / ou la prédiction d'évolution et / ou la prise en charge des troubles ou des symptômes. Nous soutenons que cette détermination pourrait être réalisée en s'appuyant sur des approches computationnelles (Huys et al., 2016, 2021; Wang & Krystal, 2014). De telles approches computationnelles s'opposent à celles des neurosciences non-computationnelles, qui se fondent majoritairement sur les catégories nosologiques du DSM, sur les construits dimensionnels (à l'image de la HiTOP) ou sur des marqueurs tels qu'identifiés au sein du projet RDoC. En d'autres termes, nous allons

¹²⁷ Une telle indépendance entre niveaux permettait également de comprendre que le vol de l'oiseau (ou tout comportement humain) correspond à un phénomène particulier : l'émergence. Un phénomène émergent signifie que la totalité des caractéristiques de l'oiseau (ou des neurones) provoque des effets qui ne sont pas visibles par l'observation d'une unique caractéristique (ou unique neurone), ou même de l'ensemble des caractéristiques agrégées (ou des neurones agrégés).

soutenir que la conception d'une variable latente peut être particulièrement utile sous certaines conditions et pour certains troubles, mais que sa détermination doit reposer sur une méthodologie computationnelle prenant en compte l'hétérogénéité individuelle, le contexte et la dynamique évolutive¹²⁸. Comme l'illustre la **Figure 1-Annexe**, les neurosciences ont tenté d'effectuer ce travail d'identification permettant de statuer sur la haute ou basse essentialité des ensembles de symptômes en psychiatrie. En d'autres termes, certains symptômes seraient plus fortement reliés que d'autres à une essence décelable. Bien qu'il soit évident que toute symptôme soit sous-tendue par (au moins) un mécanisme (voir annexe précédente), nous proposons que *la possibilité* d'identifier un tel mécanisme doit être présentée sur un *continuum* (que cette *possibilité est graduelle*). On retrouve dans cette conception une forme de pluralisme non-intégratif saupoudré (Kendler, 2005; Schaffner, 2000), dans lequel certains niveaux (de la génétique à l'environnement en passant par le niveau psychologique) seraient plus informatifs et pertinents pour un trouble donné (et par extension, pour un symptôme donné). À la suite de Hitchcock (2021), nous soutenons donc qu'il est pertinent de graduer le statut d'essentialité des symptômes.

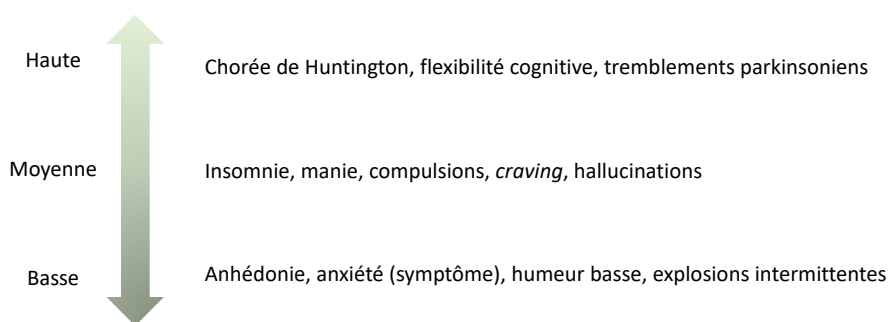


Figure 1-Annexe. Estimations du degré d'essentialité potentiel (haut, moyen ou bas) pour certains symptômes psychiatriques. Cette essentialité (estimée à l'heure actuelle et sans base empirique probante) doit être adaptée en fonction des progrès scientifiques.

Par exemple, il existe des biais cognitifs qui augmentent la focalisation sur les informations négatives. La présence de ces biais est impliquée en tant que facteur de risque pour divers troubles de l'intériorisation. Cependant, la raison pour laquelle un individu développe un trouble obsessionnel-compulsif, tandis qu'un autre développe un trouble dépressif caractérisé, n'est pas expressément connue (face à des événements contextuels et environnementaux similaires et malgré l'absence d'autres facteurs cognitifs). Cette méconnaissance des causes de différences rend difficile la possibilité d'identification d'une essence (unique ou multiple) en psychiatrie. Dans ce cadre, le trouble obsessionnel-compulsif est un trouble à basse essentialité potentielle. Nous allons décrire ce qu'est la « distance causale » au sein des troubles à basse et haute essentialité potentielle. Un trouble à basse essentialité potentielle présente une distance causale importante entre la cause et la manifestation. Ces facteurs causaux distaux s'expriment en termes biologiques, ou de temporalité, par exemple dans le cas de victimes de maltraitance dans l'enfance survenues à distance de l'événement. Ils peuvent être liés à des fluctuations de la propension à ruminer ou des

¹²⁸ Fried proposait une feuille de route pour envisager les troubles d'une nouvelle manière (Fried, 2021) : « Tout d'abord, [...] nous devrions envisager de porter notre attention sur des éléments qui vont au-delà des marqueurs biologiques ou des symptômes isolés. Deuxièmement, nous devrions étudier les systèmes dans leurs écologies respectives [...]. Troisièmement, il faudrait comprendre les troubles mentaux comme des entités dynamiques qui ne peuvent plus être décomposées en de simples relations de cause à effet. [...]. Quatrièmement, les systèmes complexes se comportent souvent de manière non intuitive : de petits changements d'entrée peuvent changer radicalement le macro-comportement, de sortie : [...] construire des modèles computationnels à l'aide d'un langage mathématique permet une interrogation plus approfondie de nos théories, par rapport aux théories narratives qui restent souvent ambiguës » (Borsboom et al., 2021).

intolérances à l'ambiguïté, qui rendent difficile la corrélation entre cause et symptômes. Nous verrons par la suite que cette distance causale doit prendre en compte les facteurs temporels, mais également contextuels (Nolen-Hoeksema et Watkins, 2011)¹²⁹. Des conditions comme les regroupements de symptômes décrites dans la schizophrénie impliquent des facteurs de risque distaux, comme une série complexe de facteurs de risque génétiques et moléculaires (McCutcheon et al., 2020), dont l'expression et les conséquences sont médiées par des facteurs tels que la maltraitance infantile, le stress familial ou la marginalisation sociale (Egerton et al., 2017). Pour les problèmes d'essentialité élevée, il existe un chemin relativement plus direct entre les mécanismes neurobiologiques et les symptômes qui en résultent.

Nous relevons trois implications à différencier ces spectres d'essentialité potentielle. Premièrement, les troubles de haute essentialité potentielle renvoient plus souvent à un mécanisme neurobiologique, tandis que ceux considérées avec une faible essentialité potentielle seraient mieux décrites selon un modèle relationnel, comme celui proposé par l'approche réseau. Deuxièmement, le modèle médical serait plus à même de décrire certains symptômes de haute essentialité tandis que le modèle de l'approche réseau (un modèle « psychologique ») semblerait plus utile pour rendre compte des symptômes de faible essentialité¹³⁰. Troisièmement, le recours à une variable latente serait plus utile pour les symptômes de haute essentialité. Il est important de prendre en compte le fait que cette approche est graduelle : en fonction de l'augmentation d'essentialité d'un groupe de symptômes corrélés, la variable latente serait de plus en plus féconde (en termes heuristiques) pour expliquer d'abord une manifestation (la voie mésocorticale dans les hallucinations), puis plusieurs à la fois (la dopamine dans les hallucinations et les idées délirantes), voire le trouble dans son intégralité (les transmissions neurochimiques cérébrales dans la schizophrénie). En dehors de quelques cas particuliers de maladies monogénétiques, nous remarquons que lorsque la variable latente (le mécanisme, en l'occurrence) explique plus qu'une unique manifestation, elle sera empiriquement plus vague, moins précise.

¹²⁹ Nous allons voir par la suite que dans les troubles d'essentialité plus basse, ce sont les interrelations dynamiques entre éléments et avec le contexte, qui fonctionnent à différentes échelles de temps, qui constituent en partie le problème (McNally, 2021).

¹³⁰ Nous utilisons ici les termes de « modèle médical » et de « modèle psychologique » de manière détournée, puisque ceux-ci s'adressent à un symptôme particulier et non à un trouble, une catégorie ou une maladie.

Bibliographie

DSM

- Akram, F., & Giordano, J. (2017). Research Domain Criteria as Psychiatric Nosology. *Cambridge Quarterly of Healthcare Ethics: CQ: The International Journal of Healthcare Ethics Committees*, 26(4), 592–601. <https://doi.org/10.1017/S096318011700010X>
- American Psychiatric Association. (1980). *American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-III®)*.
- Andreasen, N. C. (2007). DSM and the Death of Phenomenology in America: An Example of Unintended Consequences. *Schizophrenia Bulletin*, 33(1), 108–112. <https://doi.org/10.1093/schbul/sbl054>
- Aragona, M. (2013a). Neopositivism and the DSM psychiatric classification. An epistemological history. Part 1: Theoretical comparison. *History of Psychiatry*, 24(2), 166–179. <https://doi.org/10.1177/0957154X12450142>
- Aragona, M. (2013b). Neopositivism and the DSM psychiatric classification. An epistemological history. Part 2: Historical pathways, epistemological developments and present-day needs: *History of Psychiatry*. <https://doi.org/10.1177/0957154X13499671>
- Bayer, R., & Spitzer, R. L. (1985). Neurosis, Psychodynamics, and DSM-III: A History of the Controversy. *Archives of General Psychiatry*, 42(2), 187–196. <https://doi.org/10.1001/archpsyc.1985.01790250081011>
- Beals, J., Novins, D. K., Spicer, P., Orton, H. D., Mitchell, C. M., Barón, A. E., & Manson, S. M. (2004). Challenges in Operationalizing the DSM-IV Clinical Significance Criterion. *Archives of General Psychiatry*, 61(12), 1197–1207. <https://doi.org/10.1001/archpsyc.61.12.1197>
- Black, D. W., & Grant, J. E. (2014). *DSM-5® Guidebook: The Essential Companion to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*. American Psychiatric Pub.
- Blashfield. (1993). Variants of Categorical and Dimensional Models. *Psychological Inquiry*, 4(2), 95–98. https://doi.org/10.1207/s15327965pli0402_3
- Blashfield, Blum, N., & Pfohl, B. (1992). The effects of changing axis II diagnostic criteria. *Comprehensive Psychiatry*, 33(4), 245–252. [https://doi.org/10.1016/0010-440X\(92\)90048-U](https://doi.org/10.1016/0010-440X(92)90048-U)
- Boissier de Sauvages de Lacroix. (1763). *Nosologia methodica sistens morborum classes, genera et species, juxtà sydenhami mentem & botanicorum ordinem*. sumptibus fratrum de Tourneis.
- Bridgman, L. (1927). The logic of modern physics. *Macmillian Company*.
- Cade, J. F. J. (1949). Lithium salts in the treatment of psychotic excitement. *Medical Journal of Australia*, 2, 349–352.
- Cahalan, S. (2019). *The Great Pretender: The Undercover Mission That Changed Our Understanding of Madness*. Grand Central Publishing.
- Chang, H. (2021). Operationalism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2021). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/fall2021/entries/operationalism/>
- Clarke, D. E., Narrow, W. E., Regier, D. A., Kuramoto, S. J., Kupfer, D. J., Kuhl, E. A., Greiner, L., & Kraemer, H. C. (2013). DSM-5 Field Trials in the United States and Canada, Part I: Study Design, Sampling Strategy, Implementation, and Analytic Approaches. *American Journal of Psychiatry*, 170(1), 43–58. <https://doi.org/10.1176/appi.ajp.2012.12070998>
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, 94, S95–

S120.

- Cooper, R. V., & Blashfield, R. (2018). The Myth of Hempel and the DSM-III. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 70, 10–19.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. <https://doi.org/10.1037/h0040957>
- Decker, H. S. (2013). *The Making of DSM-III: A Diagnostic Manual's Conquest of American Psychiatry*. OUP USA.
- Demazeux, S. (2013a). L'échec du DSM-5, ou la victoire du principe de conservatisme. *L'information psychiatrique, Volume 89*(4), 295–302.
- Demazeux, S. (2013b). DSM-5: Pourquoi avoir supprimé le système multiaxial ? *European Psychiatry*, 28(S2), 34–35. <https://doi.org/10.1016/j.eurpsy.2013.09.084>
- Demazeux, S., & Singy, P. (Eds.). (2015). *The DSM-5 in Perspective: Philosophical Reflections on the Psychiatric Babel*. Springer Netherlands. <https://doi.org/10.1007/978-94-017-9765-8>
- Ellenberger, H. F. (1981). *The Discovery Of The Unconscious: The History And Evolution Of Dynamic Psychiatry* (Clean&Tight Contents édition). Basic Books.
- Engstrom, & Kendler, K. S. (2015). Emil Kraepelin: Icon and Reality. *American Journal of Psychiatry*, 172(12), 1190–1196. <https://doi.org/10.1176/appi.ajp.2015.15050665>
- Eysenck, H. J., Wakefield, J. A., & Friedman, A. F. (1983). Diagnosis and Clinical Assessment: The DSM-III. *Annual Review of Psychology*, 34(1), 167–193. <https://doi.org/10.1146/annurev.ps.34.020183.001123>
- Fagot-Largeault, A. (1991). Les Causes de la Mort. Histoire Naturelle Et Facteurs de Risque. *Revue de Métaphysique et de Morale*, 96(1), 97–103.
- Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., & Munoz, R. (1972). Diagnostic Criteria for Use in Psychiatric Research. *Archives of General Psychiatry*, 26(1), 57–63. <https://doi.org/10.1001/archpsyc.1972.01750190059011>
- First, M. (2004). Dimensional Models of Personality Disorders: Etiology, pathology, phenomenology, and treatment. The future of psychiatric diagnosis: Refining the research agenda. *American Psychiatric Publishing*.
- Garrabé, J. (2011). Approche historique des classifications en psychiatrie. *Annales Médico-Psychologiques, Revue Psychiatrique*. <https://doi.org/10.1016/j.amp.2011.03.002>
- Garud, R., Kumaraswamy, A., & Karnøe, P. (2010). Path Dependence or Path Creation? *Journal of Management Studies*, 47(4), 760–774. <https://doi.org/10.1111/j.1467-6486.2009.00914.x>
- Ghaemi, S. N. (2004). *The Concepts of Psychiatry: A Pluralistic Approach to the Mind and Mental Illness*. JHU Press.
- Gibson, J. G. (1973). *Psychiatric Diagnosis In New York And London* By J. E. Cooper, R. E. Kendell, B. J. Gurland, L. Sharpe, J. R. M. Copeland, and R. Simon. (Pp. 152; 5 figures, £3.) Oxford University Press: London. 1972. *Psychological Medicine*, 3(1), 130–131. <https://doi.org/10.1017/S0033291700046444>
- Goodwin, D. W., Powell, B., Bremer, D., Hoine, H., & Stern, J. (1969). Alcohol and Recall: State-Dependent Effects in Man. *Science*. <https://www.science.org/doi/abs/10.1126/science.163.3873.1358>
- Grob, G. (Ed.). (1985). *The Inner World of American Psychiatry 1890-1940: Selected Correspondence* (1st Edition). Rutgers Univ Pr.
- Guze, S. B. (1978). Nature of psychiatric illness: Why psychiatry is a branch of medicine. *Comprehensive Psychiatry*, 19(4), 295–307. [https://doi.org/10.1016/0010-440X\(78\)90012-3](https://doi.org/10.1016/0010-440X(78)90012-3)
- Hacking, I. (2013, August 8). Lost in the Forest [Review of *DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*]. *London Review of Books*, 35(15). <https://www.lrb.co.uk/the-paper/v35/n15/ian-hacking/lost-in-the-forest>
- Heckers, S., & Kendler, K. S. (2020). The evolution of Kraepelin's nosological principles. *World Psychiatry*, 19(3), 381–388. <https://doi.org/10.1002/wps.20774>
- Hempel. (1965). *Aspects of Scientific Explanation*. Free Press.
- Henderson, D. (1950). *A Text-Book of Psychiatry for Students and Practitioners* (seventh edition). Oxford University Press.
- Hess, V., & Majerus, B. (2011). Writing the history of psychiatry in the 20th century. *History of Psychiatry*, 22(2), 139–145. <https://doi.org/10.1177/0957154X11404791>
- Hoff, P. (2015). The Kraepelinian tradition. *Dialogues in Clinical Neuroscience*, 17(1), 31–41.
- Horwitz, A. V. (2014). DSM - I and DSM - II. In *The Encyclopedia of Clinical Psychology* (pp. 1–6). American

- Cancer Society. <https://doi.org/10.1002/9781118625392.wbecp012>
- Horwitz, A. V., & Wakefield, J. C. (2007). *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder*. Oxford University Press.
- Houts, A. C. (2001). The diagnostic and statistical manual's new white coat and circularity of plausible dysfunctions: Response to Wakefield, part 1. *Behaviour Research and Therapy*, 39(3), 315–345. [https://doi.org/10.1016/s0005-7967\(00\)00069-3](https://doi.org/10.1016/s0005-7967(00)00069-3)
- Jablensky, & Kendell. (2002). *Criteria for Assessing a Classification in Psychiatry*. 1–24. <https://doi.org/10.1002/047084647X.ch1>
- Johnstone, E., Cunningham-Owens, D. G., Lawrie, S., McIntosh, A., & Sharpe, M. D. (2010). *Companion to Psychiatric Studies*. Elsevier Health Sciences.
- Jones. (1944). The diagnostic of rheumatic fever. *Journal of the American Medical Association*, 126(8), 481–484. <https://doi.org/10.1001/jama.1944.02850430015005>
- Kawa, S., & Giordano, J. (2012). A brief historicity of the Diagnostic and Statistical Manual of Mental Disorders: Issues and implications for the future of psychiatric canon and practice. *Philosophy, Ethics, and Humanities in Medicine : PEHM*, 7, 2. <https://doi.org/10.1186/1747-5341-7-2>
- Kendler. (2009). An historical framework for psychiatric nosology. *Psychological Medicine*, 39(12), 1935–1941. <https://doi.org/10.1017/S0033291709005753>
- Kendler, & Engstrom, E. J. (2018). Criticisms of Kraepelin's Psychiatric Nosology: 1896–1927. *American Journal of Psychiatry*, 175(4). <https://doi.org/10.1176/appi.ajp.2017.17070730>
- Kendler, K. S. (2013). A history of the DSM-5 scientific review committee. *Psychological Medicine*, 43(9), 1793–1800. <https://doi.org/10.1017/S0033291713001578>
- Kendler, K. S. (2018). Classification of psychopathology: Conceptual and historical background. *World Psychiatry*, 17(3), 241–242. <https://doi.org/10.1002/wps.20549>
- Kendler, K. S., Muñoz, R. A., & Murphy, G. (2010). The Development of the Feighner Criteria: A Historical Perspective. *American Journal of Psychiatry*, 167(2), 134–142. <https://doi.org/10.1176/appi.ajp.2009.09081155>
- Kendler, & Parnas, J. (2012). *Philosophical Issues in Psychiatry II: Nosology*. OUP Oxford.
- Kendler, & Parnas, J. (Eds.). (2017). *Philosophical Issues in Psychiatry IV: Psychiatric Nosology*. Oxford University Press.
- Keuck, L. (2011). How to Make Sense of Broadly Applied Medical Classification Systems: Introducing Epistemic Hubs. *History and Philosophy of the Life Sciences*, 33, 583–601.
- Kirk, S. A., & Kutchins, H. (1992). *The selling of DSM: The rhetoric of science in psychiatry* (pp. x, 270). Aldine de Gruyter.
- Klerman, G. L. (1986). The National Institute of Mental Health—Epidemiologic Catchment Area (NIMH-ECA) program. *Social Psychiatry*, 21(4), 159–166. <https://doi.org/10.1007/BF00583995>
- Kupfer, D. (2005). Dimensional Models for Research and Diagnosis: A Current Dilemma. *Journal of Abnormal Psychology*, 114(4).
- Leonhard, K. (1999). *Classification of Endogenous Psychoses and their Differentiated Etiology*. Springer Science & Business Media.
- Liebowitz, S. J., & Margolis, S. E. (1990). The Fable of the Keys. *The Journal of Law & Economics*, 33(1), 1–25.
- Mellor, C. S. (1970). First Rank Symptoms of Schizophrenia: I. the Frequency in Schizophrenics on Admission to Hospital II. Differences between Individual First Rank Symptoms. *The British Journal of Psychiatry*, 117(536), 15–23. <https://doi.org/10.1192/S0007125000192116>
- Merton, R. K. (1968). The Matthew Effect in Science: The reward and communication systems of science are considered. *Science*, 159(3810), 56–63. <https://doi.org/10.1126/science.159.3810.56>
- Paris, J., & Philipps, J. (2013). *Making the DSM-5* (Vol. 2). <https://link.springer.com/book/10.1007/978-1-4614-6504-1>
- Pressman, J. D. (1997). Psychiatry and Its Origins. *Bulletin of the History of Medicine*, 71(1), 129–139.
- Purtell, J. J., Robins, E., & Cohen, M. E. (1951). Observations on clinical aspects of hysteria: A quantitative study of 50 hysteria patients and 156 control subjects. *Journal of the American Medical Association*, 146(10), 902–909. <https://doi.org/10.1001/jama.1951.03670100022006>
- Quang, G. L. (2015). L'interne devant les théories. De l'articulation entre théorie et pratique en psychiatrie. *L'information psychiatrique, Volume 91*(2), 177–182.
- Robins, E., & Mensh, I. N. (1954). Prediction in the clinical method and interrelations of biochemistry,

- psychiatry, and psychology. *The Journal of Abnormal and Social Psychology*, 49(3), 435–442. <https://doi.org/10.1037/h0056351>
- Robins, J.K., Wittchen, H.-U., Helzer, J. E., Babor, T. F., Burke, J., Farmer, A., Jablenski, A., Pickens, R., & Regier, D. (1989). The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry*, 45, 1069–1077.
- Romelli, K., Frigerio, A., & Colombo, M. (2016). DSM over time: From legitimisation of authority to hegemony. *BioSocieties*, 11(1), 1–21. <https://doi.org/10.1057/biosoc.2015.1>
- Ropes, M. W., Bennett, G. A., Cobb, S., Jacox, R., & Jessar, R. A. (1957). Proposed Diagnostic Criteria for Rheumatoid Arthritis *. *Annals of the Rheumatic Diseases*, 16(1), 118–125.
- Rosenhan, D. L. (1973). On Being Sane in Insane Places. *Science*. <https://www.science.org/doi/abs/10.1126/science.179.4070.250>
- Ruphy, S. (2016). Scientific pluralism reconsidered: A new approach to the (dis)unity of science. In *Scientific Pluralism Reconsidered: A New Approach to the (Dis)Unity of Science* (p. 186).
- Sadegh-Zadeh, K. (2008). The Prototype Resemblance Theory of Disease. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, 33(2), 106–139. <https://doi.org/10.1093/jmp/jhn004>
- Sadler, J. Z. (2005). *Values and Psychiatric Diagnosis*. OUP Oxford.
- Sandford, S. (2017). Freud, Bion and Kant: Epistemology and anthropology in The Interpretation of Dreams. *The International Journal of Psychoanalysis*, 98(1), 91–110. <https://doi.org/10.1111/1745-8315.12564>
- Scheff, T. J. (1971). *Being Mentally Ill: A Sociological Theory*. Transaction Publishers.
- Schneider, K. (1959). *Clinical psychopathology*. (Trans. By M. W. Hamilton), 5th ed (pp. xvi, 173). Grune & Stratton.
- Shepherd, J. C. (1987). *The Clinical Roots of the Schizophrenia Concept: Translations of Seminal European Contributions on Schizophrenia*. CUP Archive.
- Shorter, E. (2008). *Before Prozac: The Troubled History of Mood Disorders in Psychiatry* (1 edition). Oxford University Press.
- Sneath, P. H. A., & Sokal, R. R. (1973). Numerical taxonomy. The principles and practice of numerical classification. *Numerical Taxonomy. The Principles and Practice of Numerical Classification*. <https://www.cabdirect.org/cabdirect/abstract/19730310919>
- Spitzer, Endicott, J., & Robins, E. (1978). Research Diagnostic Criteria: Rationale and Reliability. *Archives of General Psychiatry*, 35(6), 773–782. <https://doi.org/10.1001/archpsyc.1978.01770300115013>
- Spitzer, R. L. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry*, 24(5), 399–411. [https://doi.org/10.1016/0010-440X\(83\)90032-9](https://doi.org/10.1016/0010-440X(83)90032-9)
- Spitzer, R. L., Endicott, J., & Robins, E. (1975). Clinical criteria for psychiatric diagnosis and DSM-III. *The American Journal of Psychiatry*, 132(11), 1187–1192. <https://doi.org/10.1176/ajp.132.11.1187>
- Spitzer, R. L., Williams, J. B., & Skodol, A. E. (1980). DSM-III: The major achievements and an overview. *The American Journal of Psychiatry*, 137(2), 151–164. <https://doi.org/10.1176/ajp.137.2.151>
- Srole, L., & Fischer, A. K. (1978). *Mental health in the metropolis: The midtown Manhattan study* (pp. xiii, 553). New York U Press.
- Standley, C. C., & Lin Tsong-Yi. (1963). Cahiers de santé publique, ISSN 0374-1591. La Place de l'épidémiologie en psychiatrie. In *Cahiers de santé publique*. Organisation mondiale de la santé.
- Stein, D. J. (2014). An integrative approach to psychiatric diagnosis and research. *World Psychiatry*, 13(1), 51–53. <https://doi.org/10.1002/wps.20104>
- Stengel, E. (1959). Classification of mental disorders. *Bulletin of the World Health Organization*, 21, 601–663.
- Szasz, T. S. (1960). The myth of mental illness. *American Psychologist*, 15(2), 113–118. <https://doi.org/10.1037/h0046535>
- Üstün, T. B., & Sartorius, N. (1995). *Mental Illness in General Health Care: An International Study*. John Wiley & Sons.
- Wakefield, J. C. (2001). The myth of DSM's invention of new categories of disorder: Houts's diagnostic discontinuity thesis disconfirmed. *Behaviour Research and Therapy*, 39(5), 575–624. [https://doi.org/10.1016/S0005-7967\(00\)00071-1](https://doi.org/10.1016/S0005-7967(00)00071-1)
- Werry, J. S. (1985). ICD 9 & DSM III CLASSIFICATION FOR THE CLINICIAN. *Journal of Child Psychology and*

- Psychiatry*, 26(1), 1–6. <https://doi.org/10.1111/j.1469-7610.1985.tb01624.x>
- Whooley, O., & Horwitz, A. V. (2013). The Paradox of Professional Success: Grand Ambition, Furious Resistance, and the Derailment of the DSM-5 Revision Process. In J. Paris & J. Phillips (Eds.), *Making the DSM-5: Concepts and Controversies* (pp. 75–92). Springer. https://doi.org/10.1007/978-1-4614-6504-1_6
- Widiger, T. A., & Clark, L. A. (2000). Toward DSM-V and the classification of psychopathology. *Psychological Bulletin*, 126(6), 946–963. <https://doi.org/10.1037/0033-2909.126.6.946>
- Wilson, M. (1993). DSM-III and the transformation of American psychiatry: A history. *The American Journal of Psychiatry*, 150(3), 399–410. <https://doi.org/10.1176/ajp.150.3.399>
- Wing, J. K. (1961). A Simple and Reliable Subclassification of Chronic Schizophrenia. *Journal of Mental Science*, 107(450), 862–875. <https://doi.org/10.1192/bjp.107.450.862>
- Wing, J. K., Cooper, J. E., & Sartorius, N. (2012). *Measurement and Classification of Psychiatric Symptoms: An Instruction Manual for the PSE and Catego Program*. Cambridge University Press.
- World Health Organization. (1990). *ICD-10: International statistical classification of diseases and related health problems: tenth revision*. World Health Organization. <https://apps.who.int/iris/handle/10665/42980>
- Yablonsky, A. I. (1980). On fundamental regularities of the distribution of scientific productivity. *Scientometrics*, 2(1), 3–34. <https://doi.org/10.1007/BF02016597>
- Zachar, P., Stoyanov, D. S., Aragona, M., & Jablensky, A. (2014). *Alternative perspectives on psychiatric validation: DSM, ICD, RDoC, and Beyond*. OUP Oxford.

RDoC

- Alexopoulos, G. S., & Arean, P. (2014). A model for streamlining psychotherapy in the RDoC era: The example of ‘Engage.’ *Molecular Psychiatry*, 19(1), 14–19. <https://doi.org/10.1038/mp.2013.150>
- Ankeny, R. A., & Leonelli, S. (2016). Repertoires: A post-Kuhnian perspective on scientific change and collaborative research. *Studies in History and Philosophy of Science Part A*, 60, 18–28. <https://doi.org/10.1016/j.shpsa.2016.08.003>
- Arens, K. (1996). Wilhelm Griesinger: Psychiatry between Philosophy and Praxis. *Philosophy, Psychiatry, & Psychology*, 3(3), 147–163. <https://doi.org/10.1353/ppp.1996.0025>
- Ayearst, L., Flett, G., & Hewitt, P. (2012). Where Is Multidimensional Perfectionism in DSM-5? A Question Posed to the DSM-5 Personality and Personality Disorders Work Group. *Personality Disorders*, 3. <https://doi.org/10.1037/a0026354>
- Beauchaine, T. P., & Hinshaw, S. P. (2020). RDoC and Psychopathology among Youth: Misplaced Assumptions and an Agenda for Future Research. *Journal of Clinical Child & Adolescent Psychology*, 49(3), 322–340. <https://doi.org/10.1080/15374416.2020.1750022>
- Bechtel, W. (2008). Mechanisms in Cognitive Psychology: What Are the Operations? *Philosophy of Science*, 75(5), 983–994. <https://doi.org/10.1086/594540>
- Bechtel, W., & Richardson, R. C. (2010). *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. MIT Press.
- Bolton, D. (2016). We need science to be useful too. *World Psychiatry*, 15(1), 37–38. <https://doi.org/10.1002/wps.20288>
- Boorse, C. (1977). Health as a Theoretical Concept. *Philosophy of Science*, 44(4), 542–573. <https://doi.org/10.1086/288768>
- Bowker, G. C., & Star, S. L. (1999). *Sorting things out: Classification and its consequences*. MIT Press.
- Bradley, K. A. L., Case, J. A. C., Freed, R. D., Stern, E. R., & Gabbay, V. (2017). Neural correlates of RDoC reward constructs in adolescents with diverse psychiatric symptoms: A Reward Flanker Task pilot study. *Journal of Affective Disorders*, 216, 36–45. <https://doi.org/10.1016/j.jad.2016.11.042>
- Brigandt, I. (2010). Beyond Reduction and Pluralism: Toward an Epistemology of Explanatory Integration

- in Biology. *Erkenntnis* (1975-), 73(3), 295–311.
- Brigandt, I. (2013a). Integration in biology: Philosophical perspectives on the dynamics of interdisciplinarity. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44(4 Pt A), 461–465. <https://doi.org/10.1016/j.shpsc.2013.09.009>
- Brigandt, I. (2013b). Systems biology and the integration of mechanistic explanation and mathematical explanation. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 44(4, Part A), 477–492. <https://doi.org/10.1016/j.shpsc.2013.06.002>
- Callaway, E. (2011). Schizophrenia “in a dish.” *Nature*. <https://doi.org/10.1038/news.2011.232>
- Callon, M., & Latour, B. (2013). *La science telle qu'elle se fait: Anthologie de la sociologie des sciences de langue anglaise*. La Découverte.
- Casey, B. J., Oliveri, M. E., & Insel, T. (2014). A neurodevelopmental perspective on the research domain criteria (RDoC) framework. *Biological Psychiatry*, 76(5), 350–353. <https://doi.org/10.1016/j.biopsych.2014.01.006>
- Clark, L. A., Cuthbert, B., Lewis-Fernández, R., Narrow, W. E., & Reed, G. M. (2017). Three Approaches to Understanding and Classifying Mental Disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, 18(2), 72–145. <https://doi.org/10.1177/1529100617727266>
- Clarke, D. E., Narrow, W. E., Regier, D. A., Kuramoto, S. J., Kupfer, D. J., Kuhl, E. A., Greiner, L., & Kraemer, H. C. (2013). DSM-5 Field Trials in the United States and Canada, Part I: Study Design, Sampling Strategy, Implementation, and Analytic Approaches. *American Journal of Psychiatry*, 170(1), 43–58. <https://doi.org/10.1176/appi.ajp.2012.12070998>
- Codet, H. (1926). *Psychiatrie*. G. Doin & cie.
- Comte, A. (1975). *Auguste Comte and Positivism: The Essential Writings*. Transaction Publishers.
- Cooper, (2005). *Classifying Madness: A Philosophical Examination of the Diagnostic and Statistical Manual of Mental Disorders*. Springer Netherlands. <https://doi.org/10.1007/1-4020-3345-1>
- Cooper, R. V. (2013). Avoiding false positives: Zones of rarity, the threshold problem, and the DSM clinical significance criterion. *Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie*, 58(11), 606–611. <https://doi.org/10.1177/070674371305801105>
- Craver, C. F. (2007). *Explaining the Brain*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199299317.001.0001>
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. <https://doi.org/10.1037/h0040957>
- Cuthbert. (2014). The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)*, 13(1), 28–35. <https://doi.org/10.1002/wps.20087>
- Cuthbert. (2019). The PRISM project: Social withdrawal from an RDoC perspective. *Neuroscience & Biobehavioral Reviews*, 97, 34–37. <https://doi.org/10.1016/j.neubiorev.2018.08.006>
- Cuthbert, B. N. (2020). The role of RDoC in future classification of mental disorders. *Dialogues in Clinical Neuroscience*, 22(1), 81–85. <https://doi.org/10.31887/DCNS.2020.22.1/bcuthbert>
- Cuthbert, & Insel, T. (2010). The Data of Diagnosis: New Approaches to Psychiatric Classification. *Psychiatry: Interpersonal and Biological Processes*, 73(4), 311–314. <https://doi.org/10.1521/psyc.2010.73.4.311>
- Cuthbert, & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine*, 11(1), 126. <https://doi.org/10.1186/1741-7015-11-126>
- Cuthbert, & Kozak. (2013). Constructing constructs for psychopathology: The NIMH research domain criteria. *Journal of Abnormal Psychology*, 122(3), 928–937. <https://doi.org/10.1037/a0034028>
- Deacon, T. W. (1989). The neural circuitry underlying primate calls and human language. *Human Evolution*, 4(5), 367–401. <https://doi.org/10.1007/BF02436435>
- Demazeux, S. (2019). *L'Éclipse du symptôme. L'observation clinique en psychiatrie, 1800-1950*. (Ithaque).
- Drury, S., & Cuthbert, B. (2015). Advancing Pediatric Psychiatry Research: Linking Neurobiological Processes to Novel Treatment and Diagnosis Through the Research Domain Criteria (RDoC) Project. *Therapeutic Innovation & Regulatory Science*, 49(5), 643–646. <https://doi.org/10.1177/2168479015596019>
- Duffy, A. (2010). The Early Natural History of Bipolar Disorder: What We Have Learned from Longitudinal

- High-Risk Research. *The Canadian Journal of Psychiatry*, 55(8), 477–485. <https://doi.org/10.1177/070674371005500802>
- Eijndhoven, P. F. P. van, Collard, R. M., Vrijzen, J. N., Geurts, D. G. M., Arias-Vasquez, A., Schellekens, A. F. A., Munckhof, E. van den, Brolsma, S. C. A., Duyser, F. A., Bergman, M. A., Oort, J. van, Tendolkar, I., & Schene, A. H. (2021). *Measuring Integrated Novel Dimensions in Neurodevelopmental and Stress-related Mental Disorders (MIND-Set): A cross-sectional comorbidity study from an RDoC perspective* (p. 2021.06.05.21256695). <https://doi.org/10.1101/2021.06.05.21256695>
- Ellenberger, H. F. (1981). *The Discovery Of The Unconscious: The History And Evolution Of Dynamic Psychiatry* (Clean&Tight Contents édition). Basic Books.
- Filipe, A. M., Bogossian, A., Zulla, R., Nicholas, D., & Lach, L. M. (2020). Developing a Canadian framework for social determinants of health and well-being among children with neurodisabilities and their families: An ecosocial perspective. *Disability and Rehabilitation*, 0(0), 1–12. <https://doi.org/10.1080/09638288.2020.1754926>
- First, M. B., & Westen, D. (2007). Classification for clinical practice: How to make ICD and DSM better able to serve clinicians. *International Review of Psychiatry*, 19(5), 473–481. <https://doi.org/10.1080/09540260701563429>
- Fodor, J. A. (1968). *Psychological explanation: An introduction to the philosophy of psychology* (pp. xxi, 165). Crown Publishing Group/Random House.
- Forest, D. (2014). *Neurosepticisme: Les sciences du cerveau sous le scalpel de l'épistémologue*. Ithaque. <https://hal-univ-paris10.archives-ouvertes.fr/hal-01551249>
- Forest, D. (2015). *Histoire des aphasies: Une anatomie de l'expression*. Presses Universitaires de France.
- Forest, D., & Granger, B. (2015). Les neurosciences actuelles ont-elles changé les paradigmes du XIXe siècle ? *PSN, Volume 13*(2), 15–21.
- Forsman, A. K., Wahlbeck, K., Aarø, L. E., Alonso, J., Barry, M. M., Brunn, M., Cardoso, G., Cattan, M., de Girolamo, G., Eberhard-Gran, M., Evans-Lacko, S., Fiorillo, A., Hansson, L., Haro, J. M., Hazo, J.-B., Hegerl, U., Katschnig, H., Knappe, S., Luciano, M., ... also on behalf of the ROAMER Consortium. (2015). Research priorities for public mental health in Europe: Recommendations of the ROAMER project. *European Journal of Public Health*, 25(2), 249–254. <https://doi.org/10.1093/eurpub/cku232>
- Frances. (2009). Whither DSM-V? *The British Journal of Psychiatry*, 195(5), 391–392. <https://doi.org/10.1192/bjp.bp.109.073932>
- Frances, A. (2013). The past, present and future of psychiatric diagnosis. *World Psychiatry*, 12(2), 111–112. <https://doi.org/10.1002/wps.20027>
- Gauld, C., Sarah, H., Lopez, L., Pierre, P., Micoulaud-Franchi, J. A., & Royant-Parola, S. (2021). Sleep health network analysis based on questionnaire data from 35,808 subjects concerned by their sleep. *Sleep Epidemiology*, 1, 100011.
- Gauld, C. (2021). Les réseaux de symptômes en psychopathologie. UGA Éditions, Grenoble, 190p.
- Gauld, C., Fournier, P., & Micoulaud-Franchi, J.-A. (2022). Why Is Psychiatry so Intimately Linked to the Brain? *Frontiers in Psychiatry*, 13. <https://www.frontiersin.org/articles/10.3389/fpsy.2022.872957>
- Giroux, E., & Gauld, C. (2023). Les approches de personnalisation et de précision peuvent-elles être utiles à la psychiatrie ? In *Les enjeux de la psychiatrie personnalisée et de précision* (Hermann).
- Gauld C., Baillet C., Micoulaud-Franchi J.A., Kervran C., Serre F., Auriacombe M. (2023). The Centrality of Craving in Network Analysis of Five Substance Use Disorders. *Drug and Alcohol Dependence* (in press).
- Gordon, E. (2003). Integrative Neuroscience in Psychiatry: The Role of a Standardized Database. *Australasian Psychiatry*, 11(2), 156–163. <https://doi.org/10.1046/j.1039-8562.2003.00533.x>
- Grüne-Yanoff, T. (2016). Interdisciplinary success without integration. *European Journal for Philosophy of Science*, 6(3), 343–360. <https://doi.org/10.1007/s13194-016-0139-z>
- Guelfi, J.-D. (2010). L'avenir des classifications des psychoses chez l'adulte. *L'information psychiatrique*, Volume 86(2), 127–134.
- Guze, S. B. (1967). The Diagnosis of Hysteria: What Are We Trying to Do? *American Journal of Psychiatry*, 124(4), 491–498. <https://doi.org/10.1176/ajp.124.4.491>
- Guze, S. B. (1975). The validity and significance of the clinical diagnosis of hysteria (Briquet's syndrome). *The American Journal of Psychiatry*, 132(2), 138–141. <https://doi.org/10.1176/ajp.132.2.138>
- Hall, Carter, A., & Barnett, A. (2017). Disease or Developmental Disorder: Competing Perspectives on the

- Neuroscience of Addiction. *Neuroethics*, 10(1), 103–110. <https://doi.org/10.1007/s12152-017-9303-1>
- Hall, J., Trent, S., Thomas, K. L., O'Donovan, M. C., & Owen, M. J. (2015). Genetic risk for schizophrenia: Convergence on synaptic pathways involved in plasticity. *Biological Psychiatry*, 77(1), 52–58. <https://doi.org/10.1016/j.biopsych.2014.07.011>
- Harris. (2014). New classification for neurodevelopmental disorders in DSM-5. *Current Opinion in Psychiatry*, 27(2). <https://doi.org/10.1097/YCO.0000000000000042>
- Haslam, N. (2002). Kinds of Kinds: A Conceptual Taxonomy of Psychiatric Categories. *Philosophy, Psychiatry, & Psychology*, 9(3), 203–217. <https://doi.org/10.1353/ppp.2003.0043>
- Hauser, T. U., Will, G.-J., Dubois, M., & Dolan, R. J. (2018). *Developmental Computational Psychiatry*. PsyArXiv. <https://doi.org/10.31234/osf.io/85prq>
- Healy, D. (1998). *The Antidepressant Era*. Harvard University Press.
- Helzer, J. E., Kraemer, H. C., Krueger, R. F., Wittchen, H.-U., Sirovatka, P. J., & Regier, D. A. (2009). *Dimensional Approaches in Diagnostic Classification: Refining the Research Agenda for DSM-V*. American Psychiatric Pub.
- Hempel. (1965). *Aspects of Scientific Explanation*. Free Press.
- Hippius, H., & Neundörfer, G. (2003). The discovery of Alzheimer's disease. *Dialogues in Clinical Neuroscience*, 5(1), 101–108.
- Hyman, S. E. (2007). Can neuroscience be integrated into the DSM-V? *Nature Reviews. Neuroscience*, 8(9), 725–732. <https://doi.org/10.1038/nrn2218>
- Hyman, S. E. (2010). The diagnosis of mental disorders: The problem of reification. *Annual Review of Clinical Psychology*, 6, 155–179. <https://doi.org/10.1146/annurev.clinpsy.3.022806.091532>
- Insel, Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research Domain Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders. *American Journal of Psychiatry*, 167(7), 748–751. <https://doi.org/10.1176/appi.ajp.2010.09091379>
- Insel, T. R. (2014). The NIMH Research Domain Criteria (RDoC) Project: Precision Medicine for Psychiatry. *American Journal of Psychiatry*, 171(4), 395–397. <https://doi.org/10.1176/appi.ajp.2014.14020138>
- Jablensky, A., Kirkbride, J. B., & Jones, P. B. (2011). Schizophrenia: The epidemiological horizon. In *Schizophrenia*, 3rd ed (pp. 185–225). Wiley-Blackwell. <https://doi.org/10.1002/9781444327298.ch10>
- Jaspers, K. (1950). *Introduction à la philosophie* (J. Hersch, Trans.; BIBLIOTHEQUES edition). 10 X 18.
- Jaspers, K. (1963). *General Psychopathology*. JHU Press.
- Kendler. (2005). Toward a Philosophical Structure for Psychiatry. *American Journal of Psychiatry*, 162(3), 433–440. <https://doi.org/10.1176/appi.ajp.162.3.433>
- Kendler. (2009). An historical framework for psychiatric nosology. *Psychological Medicine*, 39(12), 1935–1941. <https://doi.org/10.1017/S0033291709005753>
- Kendler. (2016). The nature of psychiatric disorders. *World Psychiatry*, 15(1), 5–12. <https://doi.org/10.1002/wps.20292>
- Kendler, K. S., & Campbell, J. (2009). Interventionist causal models in psychiatry: Repositioning the mind-body problem. *Psychological Medicine*, 39(6), 881–887. <https://doi.org/10.1017/S0033291708004467>
- Kerns, J., Cohen, J., MacDonald, A., Cho, R., Stenger, V., & Carter, C. (2004). Anterior Cingulate Conflict Monitoring and Adjustments in Control. *Science (New York, N.Y.)*, 303, 1023–1026. <https://doi.org/10.1126/science.1089910>
- Kincaid, H., & Sullivan, J. A. (2014). *Classifying Psychopathology: Mental Kinds and Natural Kinds*. MIT Press.
- Kirmayer, L. J., Lemelson, R., & Cummings, C. A. (2015). *Re-Visioning Psychiatry: Cultural Phenomenology, Critical Neuroscience, and Global Mental Health*. Cambridge University Press.
- Klerman, G. L. (1977). Mental illness, the medical model, and psychiatry. *The Journal of Medicine and Philosophy*, 220–243.
- Koudys, J. W., Traynor, J. M., Rodrigo, A. H., Carcone, D., & Ruocco, A. C. (2019). The NIMH Research Domain Criteria (RDoC) Initiative and Its Implications for Research on Personality Disorder. *Current Psychiatry Reports*, 21(6), 37. <https://doi.org/10.1007/s11920-019-1023-2>
- Kozak, M. J., & Cuthbert, B. N. (2016). The NIMH Research Domain Criteria Initiative: Background, Issues, and Pragmatics. *Psychophysiology*, 53(3), 286–297. <https://doi.org/10.1111/psyp.12518>

- Krueger, R. F., & Eaton, N. R. (2015). Transdiagnostic factors of mental disorders. *World Psychiatry, 14*(1), 27–29. <https://doi.org/10.1002/wps.20175>
- Lemoine, M. (2017). Neither from words, nor from visions: Understanding p-medicine from innovative treatments. *Lato Sensu: Revue de La Société de Philosophie Des Sciences, 4*(2), 12–23. <https://doi.org/10.20416/lrsrps.v4i2.793>
- Lilienfeld, S. O., & Treadway, M. T. (2016). Clashing Diagnostic Approaches: DSM-ICD Versus RDoC. *Annual Review of Clinical Psychology, 12*(1), 435–463. <https://doi.org/10.1146/annurev-clinpsy-021815-093122>
- Maj, M. (2014). Keeping an open attitude towards the RDoC project. *World Psychiatry, 13*(1), 1–3. <https://doi.org/10.1002/wps.20111>
- Maj, M. (2020). Beyond diagnosis in psychiatric practice. *Annals of General Psychiatry, 19*(1), 27. <https://doi.org/10.1186/s12991-020-00279-2>
- Mansell, W. (2019). Transdiagnostic psychiatry goes above and beyond classification. *World Psychiatry, 18*(3), 360–361. <https://doi.org/10.1002/wps.20680>
- Mattu, J., & Sullivan, J. A. (2021). Classification, kinds, taxonomic stability and conceptual change. *Aggression and Violent Behavior, 59*, 101477. <https://doi.org/10.1016/j.avb.2020.101477>
- McCoy, L. G., Nagaraj, S., Morgado, F., Harish, V., Das, S., & Celi, L. A. (2020). What do medical students actually need to know about artificial intelligence? *Npj Digital Medicine, 3*(1), 1–3. <https://doi.org/10.1038/s41746-020-0294-7>
- McGorry, & Nelson. (2019). Transdiagnostic psychiatry: Premature closure on a crucial pathway to clinical utility for psychiatric diagnosis. *World Psychiatry, 18*(3), 359–360. <https://doi.org/10.1002/wps.20679>
- Mittal, V. A., & Wakschlag, L. S. (2017). Research Domain Criteria (RDoC) Grows Up: Strengthening Neurodevelopmental Investigation within the RDoC Framework. *Journal of Affective Disorders, 216*, 30–35. <https://doi.org/10.1016/j.jad.2016.12.011>
- Monestès, J.-L., & Baeyens, C. (2016). *L'approche transdiagnostique en psychopathologie*. Dunod.
- Murphy, D. (2006). *Psychiatry in the Scientific Image*. A Bradford Book.
- Murphy, D. (2013). The medical model and the philosophy of science. In *The Oxford handbook of philosophy and psychiatry* (pp. 966–986). Oxford University Press.
- Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Harcourt, Brace & World.
- Neumärker, K.-J., & Bartsch, A. J. (2003). Karl Kleist (1879–1960) – A Pioneer of Neuropsychiatry. *History of Psychiatry, 14*(4), 411–458. <https://doi.org/10.1177/0957154X03144001>
- Owen, M. J., O'Donovan, M. C., Thapar, A., & Craddock, N. (2011). Neurodevelopmental hypothesis of schizophrenia. *The British Journal of Psychiatry, 198*(3), 173–175. <https://doi.org/10.1192/bjp.bp.110.084384>
- Parnas, J. (2014). The RDoC program: Psychiatry without psyche? *World Psychiatry, 13*(1), 46–47. <https://doi.org/10.1002/wps.20101>
- Pijnenburg, R., Scholtens, L. H., Ardesch, D. J., de Lange, S. C., Wei, Y., & van den Heuvel, M. P. (2021). Myelo- and cytoarchitectonic microstructural and functional human cortical atlases reconstructed in common MRI space. *NeuroImage, 239*, 118274. <https://doi.org/10.1016/j.neuroimage.2021.118274>
- Posner, M. I., & Raichle, M. E. (1994). *Images of mind* (pp. ix, 257). Scientific American Library/Scientific American Books.
- Power, R. A., Parkhill, J., & de Oliveira, T. (2017). Microbial genome-wide association studies: Lessons from human GWAS. *Nature Reviews Genetics, 18*(1), 41–50. <https://doi.org/10.1038/nrg.2016.132>
- Ramachandran. (2000). Mirror neurons and imitation learning as the driving force behind “the great leap forward” in human evolution. *Edge, 7*.
- Ramstead, M. J. D., Kirchhoff, M. D., Constant, A., & Friston, K. J. (2021). Multiscale integration: Beyond internalism and externalism. *Synthese, 198*(1), 41–70. <https://doi.org/10.1007/s11229-019-02115-x>
- Riglin, L., Hammerton, G., Heron, J., Collishaw, S., Arseneault, L., Thapar, A. K., Maughan, B., O'Donovan, M. C., & Thapar, A. (2019). Developmental Contributions of Schizophrenia Risk Alleles and Childhood Peer Victimization to Early-Onset Mental Health Trajectories. *American Journal of Psychiatry, 176*(1), 36–43. <https://doi.org/10.1176/appi.ajp.2018.18010075>

- Robins, E., & Guze, S. B. (1970). Establishment of Diagnostic Validity in Psychiatric Illness: Its Application to Schizophrenia. *American Journal of Psychiatry*, 126(7), 983–987. <https://doi.org/10.1176/ajp.126.7.983>
- Rüppel, J. (2019). “Now Is a Time for Optimism”: The Politics of Personalized Medicine in Mental Health Research. *Science, Technology, & Human Values*, 44(4), 581–611. <https://doi.org/10.1177/0162243919845049>
- Salles, J., Lacassagne, E., Benvegno, G., Berthoumieu, S. Ç., Franchitto, N., & Tauber, M. (2020). The RDoC approach for translational psychiatry: Could a genetic disorder with psychiatric symptoms help fill the matrix? the example of Prader–Willi syndrome. *Translational Psychiatry*, 10(1), 1–10. <https://doi.org/10.1038/s41398-020-00964-6>
- Sanislow, C. A., Pine, D. S., Quinn, K. J., Kozak, M. J., Garvey, M. A., Heihsen, R. K., Wang, P. S.-E., & Cuthbert, B. N. (2010). Developing constructs for psychopathology research: Research domain criteria. *Journal of Abnormal Psychology*, 119(4), 631–639. <https://doi.org/10.1037/a0020909>
- Shah, J. L., Scott, J., McGorry, P. D., Cross, S. P. M., Keshavan, M. S., Nelson, B., Wood, S. J., Marwaha, S., Yung, A. R., Scott, E. M., Öngür, D., Conus, P., Henry, C., & Hickie, I. B. (2020). Transdiagnostic clinical staging in youth mental health: A first international consensus statement. *World Psychiatry*, 19(2), 233–242. <https://doi.org/10.1002/wps.20745>
- Shepherd, M. (1982). Karl Jaspers: General Psychopathology. *The British Journal of Psychiatry*, 141(3), 310–312. <https://doi.org/10.1192/S0007125000173866>
- Southgate, V., & Hamilton, A. (2008). Unbroken mirrors: Challenging a theory of Autism. *Trends in Cognitive Sciences*, 12, 225–229. <https://doi.org/10.1016/j.tics.2008.03.005>
- Spitzer, Endicott, J., & Robins, E. (1978). Research Diagnostic Criteria: Rationale and Reliability. *Archives of General Psychiatry*, 35(6), 773–782. <https://doi.org/10.1001/archpsyc.1978.01770300115013>
- Sporns, O., Tononi, G., & Kötter, R. (2005). The Human Connectome: A Structural Description of the Human Brain. *PLOS Computational Biology*, 1(4), e42. <https://doi.org/10.1371/journal.pcbi.0010042>
- Star, S. L., & Griesemer, J. R. (1989). Institutional Ecology, “Translations” and Boundary Objects: Amateurs and Professionals in Berkeley’s Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science*, 19(3), 387–420.
- Sullivan, J. A. (2017a). Long-Term Potentiation: One Kind or Many? In M. P. Adams, Z. Biener, U. Feest, & J. A. Sullivan (Eds.), *Eppur si muove: Doing History and Philosophy of Science with Peter Machamer: A Collection of Essays in Honor of Peter Machamer* (pp. 127–140). Springer International Publishing. https://doi.org/10.1007/978-3-319-52768-0_9
- Sullivan, J. A. (2017b). Coordinated pluralism as a means to facilitate integrative taxonomies of cognition. *Philosophical Explorations*, 20(2), 129–145. <https://doi.org/10.1080/13869795.2017.1312497>
- Swedo, S. E., Baird, G., Cook, E. H., Happé, F. G., Harris, J. C., Kaufmann, W. E., King, B. H., Lord, C. E., Piven, J., Rogers, S. J., Spence, S. J., Wetherby, A., & Wright, H. H. (2012). Commentary from the DSM-5 Workgroup on Neurodevelopmental Disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(4), 347–349. <https://doi.org/10.1016/j.jaac.2012.02.013>
- Tabb, K. (2020). Should Psychiatry Be Precise? Reduction, Big Data, and Nosological Revision in Mental Health Research 1. *Philosophical Issues in Psychiatry* 5. https://www.academia.edu/38775730/Should_Psychiatry_Be_Precise_Reduction_Big_Data_and_Nosological_Revision_in_Mental_Health_Research_1
- Telles-Correia, D., & Sampaio, D. (2016). Editorial: Historical Roots of Psychopathology. *Frontiers in Psychology*, 7, 905. <https://doi.org/10.3389/fpsyg.2016.00905>
- the LHA genetics group, Xu, M. K., Gaysina, D., Barnett, J. H., Scoriels, L., van de Lagemaat, L. N., Wong, A., Richards, M., Croudace, T. J., & Jones, P. B. (2015). Psychometric precision in phenotype definition is a useful step in molecular genetic investigation of psychiatric disorders. *Translational Psychiatry*, 5(6), e593–e593. <https://doi.org/10.1038/tp.2015.86>
- Tsou. (2015). DSM-5 and Psychiatry’s Second Revolution: Descriptive Vs. Theoretical Approaches to Psychiatric Classification. In S. Demazeux & P. Singy (Eds.), *The DSM-5 in Perspective: Philosophical Reflections on the Psychiatric Babel* (pp. 43–62). Springer.
- Tsou, J. Y. (2011). The Importance of History for Philosophy of Psychiatry: The Case of the DSM and Psychiatric Classification. *Journal of the Philosophy of History*, 5(3), 446–470. <https://doi.org/10.1163/187226311X599907>

- Veissière, S. P. L., Constant, A., Ramstead, M. J. D., Friston, K. J., & Kirmayer, L. J. (2020). Thinking through other minds: A variational approach to cognition and culture. *Behavioral and Brain Sciences*, 43. <https://doi.org/10.1017/S0140525X19001213>
- von Economo, & Koskinas. (1925). *Die cytoarchitektonik der hirnrinde des erwachsenen menschen*.
- Wakefield. (1992). The concept of mental disorder. On the boundary between biological facts and social values. *The American Psychologist*, 47(3), 373–388. <https://doi.org/10.1037//0003-066x.47.3.373>
- Wakefield. (2007). The concept of mental disorder: Diagnostic implications of the harmful dysfunction analysis. *World Psychiatry*, 6(3), 149–156.
- Wakefield, J. C. (2014). Wittgenstein's nightmare: Why the RDoC grid needs a conceptual dimension. *World Psychiatry*, 13(1), 38–40. <https://doi.org/10.1002/wps.20097>
- Wakefield, J. C., & First, M. B. (2013). Diagnostic Validity and the Definition of Mental Disorder: A Program for Conceptually Advancing Psychiatry. *The Canadian Journal of Psychiatry*, 58(12), 653–655. <https://doi.org/10.1177/070674371305801201>
- Weber, M. (1904). Die "Objektivität" sozialwissenschaftlicher und sozialpolitischer Erkenntnis. *Archiv für Sozialwissenschaft und Sozialpolitik*, 19(1), 22–87.
- Wernicke, C. (1874). *Der aphasische Symptomencomplex: Eine psychologische Studie auf anatomischer Basis*. Cohn.
- Whitaker, H. A., & Etlinger, S. C. (1993). Theodor Meynert's Contribution to Classical 19th Century Aphasia Studies. *Brain and Language*, 45(4), 560–571. <https://doi.org/10.1006/brln.1993.1060>
- Wittgenstein, L. (1973). *Philosophical Investigations* (G. E. M. Anscombe, Trans.; 3rd edition). Pearson.
- Wong, E. H. F., Yocca, F., Smith, M. A., & Lee, C.-M. (2010). Challenges and opportunities for drug discovery in psychiatric disorders: The drug hunters' perspective. *The International Journal of Neuropsychopharmacology*, 13(9), 1269–1284. <https://doi.org/10.1017/S1461145710000866>
- Zachar, P., Stoyanov, D. S., Aragona, M., & Jablensky, A. (2014). *Alternative perspectives on psychiatric validation: DSM, ICD, RDoC, and Beyond*. OUP Oxford.
- Zandersen, M., & Parnas, J. (2019). Identity Disturbance, Feelings of Emptiness, and the Boundaries of the Schizophrenia Spectrum. *Schizophrenia Bulletin*, 45(1), 106–113. <https://doi.org/10.1093/schbul/sbx183>

HiTOP

-
- Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs: General and Applied*, 80(7), 1–37. <https://doi.org/10.1037/h0093906>
- Achenbach, T. M., Ivanova, M. Y., & Rescorla, L. A. (2017). Empirically based assessment and taxonomy of psychopathology for ages 1½–90+ years: Developmental, multi-informant, and multicultural findings. *Comprehensive Psychiatry*, 79, 4–18. <https://doi.org/10.1016/j.comppsy.2017.03.006>
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.
- Andrews, G., Goldberg, D. P., Krueger, R. F., Carpenter, W. T., Hyman, S. E., Sachdev, P., & Pine, D. S. (2009). Exploring the feasibility of a meta-structure for DSM-V and ICD-11: Could it improve utility and validity? *Psychological Medicine*, 39(12), 1993–2000. <https://doi.org/10.1017/S0033291709990250>
- Baldwin, D. S., & Kosky, N. (2007). Off-label prescribing in psychiatric practice. *Advances in Psychiatric Treatment*, 13(6), 414–422. <https://doi.org/10.1192/apt.bp.107.004184>
- Barbut, M., & Monjardet, B. (1970). *Ordre et classification: Algebre et combinatoire*. Hachette.
- Barlow, D. H., Farchione, T. J., Bullis, J. R., Gallagher, M. W., Murray-Latin, H., Sauer-Zavala, S., Bentley, K. H., Thompson-Hollands, J., Conklin, L. R., Boswell, J. F., Ametaj, A., Carl, J. R., Boettcher, H. T., & Cassiello-Robbins, C. (2017). The Unified Protocol for Transdiagnostic Treatment of Emotional Disorders Compared With Diagnosis-Specific Protocols for Anxiety Disorders: A Randomized Clinical Trial. *JAMA Psychiatry*, 74(9), 875–884. <https://doi.org/10.1001/jamapsychiatry.2017.2164>

- Benzecri, J. P. (1975). L'analyse des données. *Population*, 30(6), 1190–1190.
- Berrios, G. E. (1996). *The History of Mental Symptoms: Descriptive Psychopathology Since the Nineteenth Century*. Cambridge University Press.
- Blashfield, R. K. (1982). Feighner et al., Invisible Colleges, and the Matthew Effect. *Schizophrenia Bulletin*, 8(1), 1–6. <https://doi.org/10.1093/schbul/8.1.1>
- Boyd, J. H., Burke, J. D., Gruenberg, E., Holzer, C. E., Rae, D. S., George, L. K., Karno, M., Stoltzman, R., McEvoy, L., & Nestadt, G. (1984). Exclusion Criteria of DSM-III: A Study of Co-occurrence of Hierarchy-Free Syndromes. *Archives of General Psychiatry*, 41(10), 983–989. <https://doi.org/10.1001/archpsyc.1984.01790210065008>
- Buffon, G.-L. L. de. (2013). *De la manière d'étudier et de traiter l'Histoire Naturelle*. Editions la Bibliothèque Digitale.
- Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M. H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T. E. (2014). The p Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders? *Clinical Psychological Science: A Journal of the Association for Psychological Science*, 2(2), 119–137. <https://doi.org/10.1177/2167702613497473>
- Clarke, D. E., Narrow, W. E., Regier, D. A., Kuramoto, S. J., Kupfer, D. J., Kuhl, E. A., Greiner, L., & Kraemer, H. C. (2013). DSM-5 Field Trials in the United States and Canada, Part I: Study Design, Sampling Strategy, Implementation, and Analytic Approaches. *American Journal of Psychiatry*, 170(1), 43–58. <https://doi.org/10.1176/appi.ajp.2012.12070998>
- Cuthbert, B. N. (2014). The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)*, 13(1), 28–35. <https://doi.org/10.1002/wps.20087>
- Demazeux, S. (2013). L'échec du DSM-5, ou la victoire du principe de conservatisme. *L'information psychiatrique, Volume 89(4)*, 295–302.
- Eaton, N. R., Krueger, R. F., Markon, K. E., Keyes, K. M., Skodol, A. E., Wall, M., Hasin, D. S., & Grant, B. F. (2013). The structure and predictive validity of the internalizing disorders. *Journal of Abnormal Psychology*, 122(1), 86–92. <https://doi.org/10.1037/a0029598>
- Embretson, S. E., & Reise, S. P. (2013). *Item Response Theory*. Psychology Press.
- Eysenck, H. J. (1944). Types of Personality: A Factorial Study of Seven Hundred Neurotics. *Journal of Mental Science*, 90(381), 851–861. <https://doi.org/10.1192/bjpp.90.381.851>
- Fagot-Largeault, A. (2010). 6. Le concept de maladie sous-jacent aux tentatives d'informatisation du diagnostic médical. *Ethique et philosophie morale*, 1, 145–172.
- Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., & Munoz, R. (1972). Diagnostic Criteria for Use in Psychiatric Research. *Archives of General Psychiatry*, 26(1), 57–63. <https://doi.org/10.1001/archpsyc.1972.01750190059011>
- First, M. (2004). Dimensional Models of Personality Disorders: Etiology, pathology, phenomenology, and treatment. The future of psychiatric diagnosis: Refining the research agenda. *American Psychiatric Publishing*.
- First, M. (2005). Mutually exclusive versus co-occurring diagnostic categories: The challenge of diagnostic comorbidity. *Psychopathology*, 38(4). <https://doi.org/10.1159/000086093>
- Forbes, M. K. (2020). *Revisions to the Hierarchical Taxonomy of Psychopathology (HiTOP) Model*. <https://osf.io/8h7m6/>
- Fulford. (2008). Values-Based Practice: A New Partner to Evidence-Based Practice and A First for Psychiatry? *Mens Sana Monographs*, 6(1), 10–21. <https://doi.org/10.4103/0973-1229.40565>
- Fulford, -, & Handa, A. (2018). Categorical and/or continuous? Learning from vascular surgery. *World Psychiatry*, 17(3), 304–305. <https://doi.org/10.1002/wps.20565>
- Ghaemi, S. N. (2018). After the failure of DSM: Clinical research on psychiatric diagnosis. *World Psychiatry*, 17(3), 301–302. <https://doi.org/10.1002/wps.20563>
- Helzer, J. E., Kraemer, H. C., Krueger, R. F., Wittchen, H.-U., Sirovatka, P. J., & Regier, D. A. (2009). *Dimensional Approaches in Diagnostic Classification: Refining the Research Agenda for DSM-V*. American Psychiatric Pub.
- Hempel, C. G. (1955). Stephen Toulmin. What kind of discipline is logic? Actes du XIème Congrès International de Philosophie, Volume V, Logique, analyse philosophique, philosophie des mathématiques, North-Holland Publishing Company, Amsterdam 1953, and Éditions E. Nauwelaerts,

- Louvain 1953, pp. 7–11. *The Journal of Symbolic Logic*, 20(1), 54–55. <https://doi.org/10.2307/2268045>
- Hopwood, C. J., Bagby, R. M., Gralnick, T., Ro, E., Ruggero, C., Mullins-Sweatt, S., Kotov, R., Bach, B., Cicero, D. C., Krueger, R. F., Patrick, C. J., Chmielewski, M., DeYoung, C. G., Docherty, A. R., Eaton, N. R., Forbush, K. T., Ivanova, M. Y., Latzman, R. D., Pincus, A. L., ... Zimmermann, J. (2020). Integrating psychotherapy with the hierarchical taxonomy of psychopathology (HiTOP). *Journal of Psychotherapy Integration*, 30(4), 477–497. <https://doi.org/10.1037/int0000156>
- Horwitz, A. V., & Wakefield, J. C. (2007). *The loss of sadness: How psychiatry transformed normal sorrow into depressive disorder* (pp. xv, 287). Oxford University Press.
- Hull, D. L. (1988). *Science as a Process an Evolutionary Account of the Social and Conceptual Development of Science*. University of Chicago Press.
- Jablensky, A. (1999). The Nature of Psychiatric Classification: Issues Beyond ICD-10 and DSM-IV. *Australian & New Zealand Journal of Psychiatry*, 33(2), 137–144. <https://doi.org/10.1046/j.1440-1614.1999.00535.x>
- Jablensky, A. (2018). The dialectic of quantity and quality in psychopathology. *World Psychiatry*, 17(3), 300–301. <https://doi.org/10.1002/wps.20562>
- Jardine, N., Jardine, P. of H. and P. of the S. N., & Sibson, R. (1971). *Mathematical Taxonomy*. Wiley.
- Kendell, R. E. (1975). *The role of diagnosis in psychiatry* (pp. viii, 176). Blackwell Scientific Publications.
- Kendler, & Myers, J. (2010). The genetic and environmental relationship between major depression and the five-factor model of personality. *Psychological Medicine*, 40(5), 801–806. <https://doi.org/10.1017/S0033291709991140>
- Keyes, K. M., Eaton, N. R., Krueger, R. F., McLaughlin, K. A., Wall, M. M., Grant, B. F., & Hasin, D. S. (2012). Childhood maltreatment and the structure of common psychiatric disorders. *The British Journal of Psychiatry*, 200(2), 107–115. <https://doi.org/10.1192/bjp.bp.111.093062>
- Klein, D. N., Kotov, R., & Bufferd, S. J. (2011). Personality and Depression: Explanatory Models and Review of the Evidence. *Annual Review of Clinical Psychology*, 7, 269–295. <https://doi.org/10.1146/annurev-clinpsy-032210-104540>
- Kotov, R., Waszczuk, M. A., Krueger, R. F., Forbes, M. K., Watson, D., Clark, L. A., Achenbach, T. M., Althoff, R. R., Ivanova, M. Y., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Moffitt, T. E., Eaton, N. R., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., ... Zimmerman, M. (2017). The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, 126(4), 454–477. <https://doi.org/10.1037/abn0000258>
- Kraemer, H., Noda, A., & O'Hara, R. (2004). Categorical versus dimensional approaches to diagnosis: Methodological challenges. *Journal of Psychiatric Research*, 38(1), 17–25. [https://doi.org/10.1016/S0022-3956\(03\)00097-9](https://doi.org/10.1016/S0022-3956(03)00097-9)
- Krueger, R. F., & Bezdjian, S. (2009). Enhancing research and treatment of mental disorders with dimensional concepts: Toward DSM-V and ICD-11. *World Psychiatry*, 8(1), 3–6.
- Krueger, R. F., Kotov, R., Watson, D., Forbes, M. K., Eaton, N. R., Ruggero, C. J., Simms, L. J., Widiger, T. A., Achenbach, T. M., Bach, B., Bagby, R. M., Bornovalova, M. A., Carpenter, W. T., Chmielewski, M., Cicero, D. C., Clark, L. A., Conway, C., DeClercq, B., DeYoung, C. G., ... Zimmermann, J. (2018). Progress in achieving quantitative classification of psychopathology. *World Psychiatry*, 17(3), 282–293. <https://doi.org/10.1002/wps.20566>
- Kutschenko, L. K. (2011). How to make sense of broadly applied medical classification systems: Introducing epistemic hubs. *History and Philosophy of the Life Sciences*, 33(4), 583–601.
- Kwako, L. E., Momenan, R., Grodin, E. N., Litten, R. Z., Koob, G. F., & Goldman, D. (2017). Addictions Neuroclinical Assessment: A reverse translational approach. *Neuropharmacology*, 122, 254–264. <https://doi.org/10.1016/j.neuropharm.2017.03.006>
- Lahey, B. B., Morre, T. M., & Kacurkin, A. N. (2021). Hierarchical models of psychopathology: Empirical support, implications, and remaining issues. *World Psychiatry*.
- Latzman, R. D., & DeYoung, C. G. (2020). Using empirically-derived dimensional phenotypes to accelerate clinical neuroscience: The Hierarchical Taxonomy of Psychopathology (HiTOP) framework. *Neuropsychopharmacology*, 45(7), 1083–1085. <https://doi.org/10.1038/s41386-020-0639-6>
- Lilienfeld, S. O., & Treadway, M. T. (2016). Clashing Diagnostic Approaches: DSM-ICD Versus RDoC. *Annual Review of Clinical Psychology*, 12(1), 435–463. <https://doi.org/10.1146/annurev-clinpsy-021815->

- Lopez, M. F., Compton, W. M., Grant, B. F., & Breiling, J. P. (2007). Dimensional approaches in diagnostic classification: A critical appraisal. *International Journal of Methods in Psychiatric Research*, *16*(S1), S6–S7. <https://doi.org/10.1002/mpr.213>
- Lorr, M., Klett, C. J., & McNair, D. M. (1963). *Syndromes of psychosis* (pp. x, 286). Macmillan Co.
- Lundahl, B., Kunz, C., Brownell, C., Tollefson, D., & Burke, B. (2010). A Meta-Analysis of Motivational Interviewing: Twenty-Five Years of Empirical Studies. *Research on Social Work Practice - RES SOCIAL WORK PRAC*, *20*, 137–160. <https://doi.org/10.1177/1049731509347850>
- Markon, K. E. (2010). How things fall apart: Understanding the nature of internalizing through its relationship with impairment. *Journal of Abnormal Psychology*, *119*(3), 447–458. <https://doi.org/10.1037/a0019707>
- Markon, K. E., Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological Bulletin*, *137*(5), 856–879. <https://doi.org/10.1037/a0023678>
- Martinez, M., Marangell, L. B., & Martinez, J. M. (2008). Psychopharmacology. In *The American Psychiatric Publishing textbook of psychiatry, 5th ed* (pp. 1053–1131). American Psychiatric Publishing, Inc.
- McGorry, P. D., & Hickie, I. B. (2019). *Clinical Staging in Psychiatry*. Cambridge University Press.
- Mohamed, S., Rosenheck, R., Swartz, M., Stroup, S., Lieberman, J. A., & Keefe, R. S. E. (2008). Relationship of Cognition and Psychopathology to Functional Impairment in Schizophrenia. *American Journal of Psychiatry*, *165*(8), 978–987. <https://doi.org/10.1176/appi.ajp.2008.07111713>
- Peirce, C. S. (1880). On the Algebra of Logic. *American Journal of Mathematics*, *3*(1), 15–57. <https://doi.org/10.2307/2369442>
- Pichot, P. (1994). Nosological models in psychiatry. *The British Journal of Psychiatry*, *164*, 232–240. <https://doi.org/10.1192/bjp.164.2.232>
- Reed, G. M. (2018). HiTOP must meet the use requirements of the ICD before it can aspire to replace it. *World Psychiatry*, *17*(3), 296–298. <https://doi.org/10.1002/wps.20560>
- Schröder, E. (1890). *Vorlesungen über die algebra der logik: Exakte logik*. B. G. Teubner.
- Simms, L., Gros, D., Watson, D., & O'Hara, M. (2008). Parsing general and specific components of depression and anxiety with bifactor modeling. *Depression and Anxiety*, *25*, E34–46. <https://doi.org/10.1002/da.20432>
- Skodol, A. E., Clark, L. A., Bender, D. S., Krueger, R. F., Morey, L. C., Verheul, R., Alarcon, R. D., Bell, C. C., Siever, L. J., & Oldham, J. M. (2011). Proposed changes in personality and personality disorder assessment and diagnosis for DSM-5 Part I: Description and rationale. *Personality Disorders: Theory, Research, and Treatment*, *2*(1), 4–22. <https://doi.org/10.1037/a0021891>
- Sneath, P. H. A., & Sokal, R. R. (1973). Numerical taxonomy. The principles and practice of numerical classification. *Numerical Taxonomy. The Principles and Practice of Numerical Classification*. <https://www.cabdirect.org/cabdirect/abstract/19730310919>
- Spitzer, Endicott, J., & Robins, E. (1978). Research Diagnostic Criteria: Rationale and Reliability. *Archives of General Psychiatry*, *35*(6), 773–782. <https://doi.org/10.1001/archpsyc.1978.01770300115013>
- Spitzer, & Wakefield. (1999). DSM-IV diagnostic criterion for clinical significance: Does it help solve the false positives problem? *The American Journal of Psychiatry*, *156*(12). <https://doi.org/10.1176/ajp.156.12.1856>
- Stephan, K. E., Bach, D. R., Fletcher, P. C., Flint, J., Frank, M. J., Friston, K. J., Heinz, A., Huys, Q. J. M., Owen, M. J., Binder, E. B., Dayan, P., Johnstone, E. C., Meyer-Lindenberg, A., Montague, P. R., Schnyder, U., Wang, X.-J., & Breakspear, M. (2016). Charting the landscape of priority problems in psychiatry, part 1: Classification and diagnosis. *The Lancet Psychiatry*, *3*(1), 77–83. [https://doi.org/10.1016/S2215-0366\(15\)00361-2](https://doi.org/10.1016/S2215-0366(15)00361-2)
- Taylor, M. J., Freeman, D., & Ronald, A. (2016). Dimensional psychotic experiences in adolescence: Evidence from a taxometric study of a community-based sample. *Psychiatry Research*, *241*, 35–42. <https://doi.org/10.1016/j.psychres.2016.04.021>
- Totah, N., Akil, H., Huys, Q., Krystal, J., MacDonald, A., Maia, T., Malenka, R., & Pauli, W. (2016). *Open Issues in Psychiatry Complexity and Heterogeneity in Psychiatric Disorders: Opportunities for Computational Psychiatry*. <https://doi.org/10.7551/mitpress/9780262035422.003.0003>
- Tyrer, P. (2018). Dimensions fit the data, but can clinicians fit the dimensions? *World Psychiatry*, *17*(3), 295–

296. <https://doi.org/10.1002/wps.20559>
- Üstün, B., & Kennedy, C. (2009). What is “functional impairment”? Disentangling disability from clinical significance. *World Psychiatry, 8*(2), 82–85.
- van Loo, H. M., Romeijn, J.-W., de Jonge, P., & Schoevers, R. A. (2013). Psychiatric comorbidity and causal disease models. *Preventive Medicine: An International Journal Devoted to Practice and Theory, 57*(6), 748–752. <https://doi.org/10.1016/j.ypmed.2012.10.018>
- van Stratten, A., Hill, J. J., Richards, D., & Cuijpers, P. (2014). *Stepped care treatment delivery for depression: A systematic review and meta-analysis*. <https://doi.org/10.1017/S0033291714000701>
- Waszczuk, M. A. (2021). The utility of hierarchical models of psychopathology in genetics and biomarker research. *World Psychiatry*.
- Waszczuk, M. A., Kotov, R., Ruggero, C., Gamez, W., & Watson, D. (2017). Hierarchical structure of emotional disorders: From individual symptoms to the spectrum. *Journal of Abnormal Psychology, 126*(5), 613–634. <https://doi.org/10.1037/abn0000264>
- Wittchen, H., & Beesdo-Baum, K. (2018). “Throwing out the baby with the bathwater”? Conceptual and methodological limitations of the HiTOP approach. *World Psychiatry, 17*(3), 298–299. <https://doi.org/10.1002/wps.20561>
- Wittenborn, J. R. (1951). Symptom patterns in a group of mental hospital patients. *Journal of Consulting Psychology, 15*(4), 290–302. <https://doi.org/10.1037/h0054947>
- World Health Organization. (2017). International Classification of Diseases 11th Edition (ICD-11). In *Encyclopedia of Personality and Individual Differences* (pp. 1–6). Springer International Publishing. https://doi.org/10.1007/978-3-319-28099-8_904-1
- Zachar, P. (2008). Real Kinds but No True Taxonomy: An Essay in Psychiatric Systematics. In K. S. Kendler & J. Parnas (Eds.), *Philosophical Issues in Psychiatry: Explanation, Phenomenology, and Nosology*. Johns Hopkins University Press.
- Zachar, P. (2018). Quantitative classification as (re-)descriptive psychopathology. *World Psychiatry, 17*(3), 294–295. <https://doi.org/10.1002/wps.20558>
- Zachar, P., & Kendler, K. S. (2007). Psychiatric disorders: A conceptual taxonomy. *The American Journal of Psychiatry, 164*(4), 557–565. <https://doi.org/10.1176/ajp.2007.164.4.557>
- Zachar, P., Krueger, R. F., & Kendler, K. S. (2016). Personality disorder in DSM-5: An oral history. *Psychological Medicine, 46*(1), 1–10. <https://doi.org/10.1017/S0033291715001543>
- Zimmerman, M., Ellison, W., Young, D., Chelminski, I., & Dalrymple, K. (2015). How many different ways do patients meet the diagnostic criteria for major depressive disorder? *Comprehensive Psychiatry, 56*, 29–34. <https://doi.org/10.1016/j.comppsy.2014.09.007>

Approche par le symptôme

- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.
- Amundson, R., & Lauder, G. V. (1994). Function without purpose. *Biology and Philosophy, 9*(4), 443–469. <https://doi.org/10.1007/BF00850375>
- Argolo, F. C., Cavalcanti-Ribeiro, P., Netto, L. R., & Quarantini, L. C. (2015). Prevention of posttraumatic stress disorder with propranolol: A meta-analytic review. *Journal of Psychosomatic Research, 79*(2), 89–93. <https://doi.org/10.1016/j.jpsychores.2015.04.006>
- Beard, C., Millner, A. J., Forgeard, M. J. C., Fried, E. I., Hsu, K. J., Treadway, M. T., et al. (2016). Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychological Medicine, 46*(16), 3359–3369. <https://doi.org/10.1017/S0033291716002300>
- Beck, A. T. (1993). Cognitive therapy: Past, present, and future. *Journal of Consulting and Clinical Psychology, 61*(2), 194–198. <https://doi.org/10.1037/0022-006X.61.2.194>
- Bertalanffy, L. von. (2015). *General System Theory: Foundations, Development, Applications*. George Braziller,

Incorporated.

- Bhugra, D., Easter, A., Mallaris, Y., & Gupta, S. (2011). Clinical decision making in psychiatry by psychiatrists. *Acta Psychiatrica Scandinavica*, 124(5), 403–411. <https://doi.org/10.1111/j.1600-0447.2011.01737.x>
- Boorse, C. (1977). Health as a Theoretical Concept. *Philosophy of Science*, 44(4), 542–573. <https://doi.org/10.1086/288768>
- Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology*, 64(9), 1089–1108. <https://doi.org/10.1002/jclp.20503>
- Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, 16(1), 5–13. <https://doi.org/10.1002/wps.20375>
- Borsboom, D., Cramer, A. O. J., & Kalis, A. (2019). Brain disorders? Not really: Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences*, 42. <https://doi.org/10.1017/S0140525X17002266>
- Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., et al. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, 1(1), 1–18. <https://doi.org/10.1038/s43586-021-00055-w>
- Borsboom, D., Fried, E. I., Epskamp, S., Waldorp, L. J., van Borkulo, C. D., van der Maas, H. L. J., & Cramer, A. O. J. (2017). False alarm? A comprehensive reanalysis of “Evidence that psychopathology symptom networks have limited replicability” by Forbes, Wright, Markon, and Krueger (2017). *Journal of Abnormal Psychology*, 126(7), 989–999. <https://doi.org/10.1037/abn0000306>
- Boyd, R. (1991). Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 61(1/2), 127–148.
- Boyd, R. (1999). Homeostasis, Species, and Higher Taxa. In R. A. Wilson (Ed.), *Species: New Interdisciplinary Essays* (pp. 141–85). MIT Press.
- Bringmann, L. F. (2021). Person-specific networks in psychopathology: Past, present, and future. *Current Opinion in Psychology*, 41, 59–64. <https://doi.org/10.1016/j.copsyc.2021.03.004>
- Bringmann, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., et al. (2019). What do centrality measures measure in psychological networks? *Journal of Abnormal Psychology*, 128(8), 892–903. <https://doi.org/10.1037/abn0000446>
- Brülde, B., & Radovic, F. (2006). What is Mental about Mental Disorder? *Philosophy, Psychiatry, & Psychology*, 13(2), 99–116. <https://doi.org/10.1353/ppp.2007.0001>
- Casali, M. (2021). Exploring the boundaries and ontology of Psychiatric Disorders (PDs) using the Homeostatic Property Cluster (HPC) model: *Lato Sensu: Revue de la Société de philosophie des sciences*, 8(2), 15–31. <https://doi.org/10.20416/LSRSPS.V8I2.3>
- Clementz, B. A., Sweeney, J. A., Hamm, J. P., Ivleva, E. I., Ethridge, L. E., Pearlson, G. D., et al. (2016). Identification of Distinct Psychosis Biotypes Using Brain-Based Biomarkers. *The American Journal of Psychiatry*, 173(4), 373–384. <https://doi.org/10.1176/appi.ajp.2015.14091200>
- Clementz, B. A., Trotti, R. L., Pearlson, G. D., Keshavan, M. S., Gershon, E. S., Keedy, S. K., et al. (2020). Testing Psychosis Phenotypes From Bipolar-Schizophrenia Network for Intermediate Phenotypes for Clinical Application: Biotype Characteristics and Targets. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, 5(8), 808–818. <https://doi.org/10.1016/j.bpsc.2020.03.011>
- Cohen, A. S., Schwartz, E., Le, T., Cowan, T., Cox, C., Tucker, R., et al. (2020). Validating digital phenotyping technologies for clinical use: the critical importance of “resolution.” *World Psychiatry*, 19(1), 114–115. <https://doi.org/10.1002/wps.20703>
- Cooper. (2002). Disease. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 33(2), 263–282. [https://doi.org/10.1016/S0039-3681\(02\)00018-3](https://doi.org/10.1016/S0039-3681(02)00018-3)
- Cooper. (2005). *Classifying Madness: A Philosophical Examination of the Diagnostic and Statistical Manual of Mental Disorders*. Springer Netherlands. <https://doi.org/10.1007/1-4020-3345-1>
- Cooper. (2012). *Is psychiatric classification a good thing? Philosophical Issues in Psychiatry II*. Oxford University Press. <https://oxfordmedicine.com/view/10.1093/med/9780199642205.001.0001/med-9780199642205-chapter-004001>. Accessed 2 September 2021
- Cooper. (2013a). Avoiding false positives: zones of rarity, the threshold problem, and the DSM clinical

- significance criterion. *Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie*, 58(11), 606–611. <https://doi.org/10.1177/070674371305801105>
- Cooper. (2013b). Natural kinds. In K. Fulford, M. Davies, G. Graham, J. Sadler, G. Stanghellini, & T. Thornton (Eds.), . Oxford: Oxford University Press. <https://eprints.lanacs.ac.uk/id/eprint/50918/>. Accessed 11 November 2021
- Cooper. (2014). *Psychiatry and Philosophy of Science*. Routledge.
- Cooper. (2020). The Concept of Disorder Revisited: Robustly Value-Laden Despite Change. *Aristotelian Society Supplementary Volume*, 94(1), 141–161. <https://doi.org/10.1093/arisup/akaa010>
- Cooper, R. (2021). On Harm. In *Defining Mental Disorder Jerome Wakefield and His Critics*. <https://direct.mit.edu/books/book/5015/chapter/2812075/On-Harm>. Accessed 21 July 2022
- Cramer, A. O. J., Borkulo, C. D. van, Giltay, E. J., Maas, H. L. J. van der, Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major Depression as a Complex Dynamic System. *PLOS ONE*, 11(12), e0167490. <https://doi.org/10.1371/journal.pone.0167490>
- Craver, C. F. (2001). Role Functions, Mechanisms, and Hierarchy. *Philosophy of Science*, 68(1), 53–74. <https://doi.org/10.1086/392866>
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. <https://doi.org/10.1037/h0040957>
- Cucuringu, M., Rombach, P., Lee, S. H., & Porter, M. A. (2016). Detection of core–periphery structure in networks using spectral methods and geodesic paths. *European Journal of Applied Mathematics*, 27(6), 846–887. <https://doi.org/10.1017/S095679251600022X>
- Cummins, R. (1975). Functional Analysis. *Journal of Philosophy*, 72(November), 741–64. <https://doi.org/10.2307/2024640>
- Cuthbert. (2014). The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World psychiatry: official journal of the World Psychiatric Association (WPA)*, 13(1), 28–35. <https://doi.org/10.1002/wps.20087>
- Darden, L. (2008). Thinking Again about Biological Mechanisms. *Philosophy of Science*, 75(5), 958–969. <https://doi.org/10.1086/594538>
- De Block, A., & Sholl, J. (2021). Harmless Dysfunctions and the Problem of Normal Variation. In *Defining Mental Disorder: Jerome Wakefield and His Critics*. <https://direct.mit.edu/books/book/5015/chapter/2812073/Harmless-Dysfunctions-and-the-Problem-of-Normal>. Accessed 23 July 2022
- De Vreese. (2021). Against the Disorder/Nondisorder Dichotomy. In *Defining Mental Disorder Jerome Wakefield and His Critics*. <https://direct.mit.edu/books/book/5015/chapter/2812050/Against-the-Disorder-Nondisorder-Dichotomy>. Accessed 21 July 2022
- De Vreese, L. (2017). How to Proceed in the Disease Concept Debate? A Pragmatic Approach. *The Journal of Medicine and Philosophy*, 42(4), 424–446. <https://doi.org/10.1093/jmp/jhx011>
- Demazeux, S. (2019). *L'Éclipse du symptôme. L'observation clinique en psychiatrie, 1800-1950*. (Ithaque.).
- Drummond, D. C., Litten, R. Z., Lowman, C., & Hunt, W. A. (2000). Craving research: Future directions. *Addiction*, 95(Suppl2), S247–S255. <https://doi.org/10.1080/09652140050111816>
- Dupré, J. (2002). *Humans and Other Animals*. <https://doi.org/10.3366/edinburgh/9780748639403.003.0006>
- Eaton, W. W., Regier, D. A., Locke, B. Z., & Taube, C. A. (1981). The Epidemiologic Catchment Area Program of the National Institute of Mental Health. *Public Health Reports*, 96(4), 319–325.
- Engelhardt, H. T. (1975). The Concepts of Health and Disease. In H. T. Engelhardt & S. F. Spicker (Eds.), *Evaluation and Explanation in the Biomedical Sciences: Proceedings of the First Trans-Disciplinary Symposium on Philosophy and Medicine Held at Galveston, May 9–11, 1974* (pp. 125–141). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-010-1769-5_9
- Engelhardt, H. T. (1981). Clinical Judgment. *Metamedicine*, 2(3), 301–317. <https://doi.org/10.1007/BF00882077>
- Engelhardt, H. T. (1984). Clinical Problems and the Concept of Disease. In L. Nordenfelt & B. I. B. Lindahl (Eds.), *Health, Disease, and Causal Explanations in Medicine* (pp. 27–41). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-009-6283-5_5
- Ereshefsky, M. (2009). Defining ‘health’ and ‘disease.’ *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 40(3), 221–227.

- <https://doi.org/10.1016/j.shpsc.2009.06.005>
- Faucher, L., & Forest, D. (Eds.). (2021). *Defining Mental Disorder: Jerome Wakefield and His Critics* (1st edition). Cambridge, Massachusetts: The MIT Press.
- Fernandes, B. S., Williams, L. M., Steiner, J., Leboyer, M., Carvalho, A. F., & Berk, M. (2017). The new field of 'precision psychiatry.' *BMC Medicine*, *15*(1), 80. <https://doi.org/10.1186/s12916-017-0849-x>
- Ferrante, M., Redish, A. D., Oquendo, M. A., Averbek, B. B., Kinnane, M. E., & Gordon, J. A. (2019). Computational psychiatry: a report from the 2017 NIMH workshop on opportunities and challenges. *Molecular Psychiatry*, *24*(4), 479–483. <https://doi.org/10.1038/s41380-018-0063-z>
- Frances, A. (2013). *Essentials of Psychiatric Diagnosis: Responding to the Challenge of DSM-5*. Guilford Publications.
- Fried, (2015). Problematic assumptions have slowed down depression research: why symptoms, not syndromes are the way forward. *Frontiers in Psychology*, *6*. <https://doi.org/10.3389/fpsyg.2015.00309>
- Fried, E. I., von Stockert, S., Haslbeck, J. M. B., Lamers, F., Schoevers, R. A., & Penninx, B. W. J. H. (2020). Using network analysis to examine links between individual depressive symptoms, inflammatory markers, and covariates. *Psychological Medicine*, *50*(16), 2682–2690. <https://doi.org/10.1017/S0033291719002770>
- Fried, Eiko I., Epskamp, S., Nesse, R. M., Tuerlinckx, F., & Borsboom, D. (2016). What are “good” depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *Journal of Affective Disorders*, *189*, 314–320. <https://doi.org/10.1016/j.jad.2015.09.005>
- Fried, Eiko I., van Borkulo, C. D., Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017a). Mental disorders as networks of problems: a review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *52*(1), 1–10. <https://doi.org/10.1007/s00127-016-1319-z>
- Fried, Eiko I., van Borkulo, C. D., Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017b). Mental disorders as networks of problems: a review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *52*(1), 1–10. <https://doi.org/10.1007/s00127-016-1319-z>
- Fusar-Poli, P., Hijazi, Z., Stahl, D., & Steyerberg, E. W. (2018). The Science of Prognosis in Psychiatry: A Review. *JAMA psychiatry*, *75*(12), 1289–1297. <https://doi.org/10.1001/jamapsychiatry.2018.2530>
- Gagné-Julien, A.-M. (2020). Towards a socially constructed and objective concept of mental disorder. *Synthese*. <https://doi.org/10.1007/s11229-020-02647-7>
- Gerin, M. I., Puetz, V. B., Blair, R. J. R., White, S., Sethi, A., Hoffmann, F., et al. (2017). A neurocomputational investigation of reinforcement-based decision making as a candidate latent vulnerability mechanism in maltreated children. *Development and Psychopathology*, *29*(5), 1689–1705. <https://doi.org/10.1017/S095457941700133X>
- Glennan, S. (2017). *The New Mechanical Philosophy*. Oxford University Press.
- Gómez-Carrillo, A., Langlois-Thérien, T., & Kirmayer, L. J. (2018). Precision Psychiatry—Yes, but Precisely What? *JAMA Psychiatry*, *75*(12), 1302–1303. <https://doi.org/10.1001/jamapsychiatry.2018.2651>
- Gross, F. (2011). What Systems Biology Can Tell Us about Disease. *History and philosophy of the life sciences*, *33*, 477–96.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *23*, 56–62. <https://doi.org/10.1136/jnnp.23.1.56>
- Haslam, N. (2002). Kinds of Kinds: A Conceptual Taxonomy of Psychiatric Categories. *Philosophy, Psychiatry, & Psychology*, *9*(3), 203–217. <https://doi.org/10.1353/ppp.2003.0043>
- Hauswald, R. (2016). The Ontology of Interactive Kinds. *Journal of Social Ontology*, *2*(2), 203–221. <https://doi.org/10.1515/jso-2015-0049>
- Herrman, H., Stewart, D. E., Diaz-Granados, N., Berger, E. L., Jackson, B., & Yuen, T. (2011). What is Resilience? *The Canadian Journal of Psychiatry*, *56*(5), 258–265. <https://doi.org/10.1177/070674371105600504>
- Hesslow, G. (1993). Do we need a concept of disease? *Theoretical Medicine*, *14*(1), 1–14. <https://doi.org/10.1007/BF00993984>
- Hitchcock, P., Fried, E. I., & Frank, M. (2021, April 1). Computational Psychiatry Needs Time and Context. PsyArXiv. <https://doi.org/10.31234/osf.io/xf2u3>
- Illari, P., & Williamson, J. (2011). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, *2*, 119–135. <https://doi.org/10.1007/s13194-011-0038-2>

- Ivleva, E. I., Clementz, B. A., Dutcher, A. M., Arnold, S. J. M., Jeon-Slaughter, H., Aslan, S., et al. (2017). Brain Structure Biomarkers in the Psychosis Biotypes: Findings From the Bipolar-Schizophrenia Network for Intermediate Phenotypes. *Biological Psychiatry*, 82(1), 26–39. <https://doi.org/10.1016/j.biopsych.2016.08.030>
- Jones, P. J., Ma, R., & McNally, R. J. (2019). Bridge Centrality: A Network Approach to Understanding Comorbidity. *Multivariate Behavioral Research*, 0(0), 1–15. <https://doi.org/10.1080/00273171.2019.1614898>
- Jones, P. J., & Robinaugh, D. (2021, June 18). So What? Implications of Network Theory for Research and Practice. PsyArXiv. <https://doi.org/10.31234/osf.io/6bqdx>
- Kendler, K. S. (2008). Explanatory models for psychiatric illness. *The American Journal of Psychiatry*, 165(6), 695–702. <https://doi.org/10.1176/appi.ajp.2008.07071061>
- Kendler, Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine*, 41(6), 1143–1150. <https://doi.org/10.1017/S0033291710001844>
- Keuck, L. (2011). How to Make Sense of Broadly Applied Medical Classification Systems: Introducing Epistemic Hubs. *History and philosophy of the life sciences*, 33, 583–601.
- Kim, N., & Ahn, W. (2002). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology General*, 131, 451–476. <https://doi.org/10.1037//0096-3445.131.4.451>
- Kincaid, H. (2008). Do We Need Theory to Study Disease? Lessons from cancer research and their implications for mental illness. *Perspectives in biology and medicine*, 51, 367–78. <https://doi.org/10.1353/pbm.0.0019>
- Kirmayer, L. J., Lemelson, R., & Cummings, C. A. (2015). *Re-Visioning Psychiatry: Cultural Phenomenology, Critical Neuroscience, and Global Mental Health*. Cambridge University Press.
- Kitano, H. (2004). Biological robustness. *Nature Reviews Genetics*, 5(11), 826–837. <https://doi.org/10.1038/nrg1471>
- Kleinman, A. (1988). *The illness narratives: Suffering, healing, and the human condition* (pp. xviii, 284). New York, NY, US: Basic Books.
- Kotov, R., Waszczuk, M. A., Krueger, R. F., Forbes, M. K., Watson, D., Clark, L. A., et al. (2017). The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of abnormal psychology*, 126(4), 454–477. <https://doi.org/10.1037/abn0000258>
- Krueger, J. (2015). Theoretical Health and Medical Practice. *Philosophy of Science*, 82(3), 491–508. <https://doi.org/10.1086/681628>
- Lahey, B. B., Flagg, E. W., Bird, H. R., Schwab-stone, M. E., Canino, G., Dulcan, M. K., et al. (1996). The NIMH Methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA) Study: Background and Methodology. *Journal of the American Academy of Child & Adolescent Psychiatry*, 35(7), 855–864. <https://doi.org/10.1097/00004583-199607000-00011>
- Lesne, A. (2008). Robustness: confronting lessons from physics and biology. *Biological Reviews*, 83(4), 509–532. <https://doi.org/10.1111/j.1469-185X.2008.00052.x>
- Lilienfeld, S. O., & Marino, L. (1995). Mental disorder as a Roschian concept: A critique of Wakefield's "harmful dysfunction" analysis. *Journal of Abnormal Psychology*, 104(3), 411–420. <https://doi.org/10.1037/0021-843X.104.3.411>
- Lux, V., & Kendler, K. S. (2010). Deconstructing major depression: a validation study of the DSM-IV symptomatic criteria. *Psychological Medicine*, 40(10), 1679–1690. <https://doi.org/10.1017/S0033291709992157>
- Ma, A., & Mondragón, R. J. (2015). Rich-Cores in Networks. *PLOS ONE*, 10(3), e0119678. <https://doi.org/10.1371/journal.pone.0119678>
- Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about Mechanisms. *Philosophy of Science*, 67(1), 1–25. <https://doi.org/10.1086/392759>
- Mackie, J. L. (1965). Causes and Conditions. *American Philosophical Quarterly*, 2(4), 245–264.
- Marková, I. S., & Berrios, G. E. (2012). Epistemology of Psychiatry. *Psychopathology*, 45(4), 220–227. <https://doi.org/10.1159/000331599>
- McGorry, Keshavan, M., Goldstone, S., Amminger, P., Allott, K., Berk, M., et al. (2014). Biomarkers and clinical staging in psychiatry. *World Psychiatry*, 13(3), 211–223. <https://doi.org/10.1002/wps.20144>
- McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy*,

- 86, 95–104. <https://doi.org/10.1016/j.brat.2016.06.006>
- Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In *Similarity and analogical reasoning* (pp. 179–195). New York, NY, US: Cambridge University Press. <https://doi.org/10.1017/CBO9780511529863.009>
- Meehl, P. (1967). What Can the Clinician Do Well? *Problems in human assessment*, pp. 594–599.
- Messick, S. (1989). Validity. In *Educational measurement, 3rd ed* (pp. 13–103). American Council on Education.
- Moghaddam-Taaheri, S. (2011). Understanding pathology in the context of physiological mechanisms: the practicality of a broken-normal view. *Biology & Philosophy*, 26(4), 603. <https://doi.org/10.1007/s10539-011-9258-2>
- Murphy, D. (2006). *Psychiatry in the Scientific Image*. Cambridge, MA, USA: A Bradford Book.
- Nelson, B., McGorry, P., & Fernandez, F. (2020). Integrating Clinical Staging and Phenomenological Psychopathology to Add Depth, Nuance, and Utility to Clinical Phenotyping: A Heuristic Challenge. Accessed 25 August 2021
- Nervi, M. (2010). Mechanisms, malfunctions and explanation in medicine. *Biology & Philosophy*, 25(2), 215–228. <https://doi.org/10.1007/s10539-009-9190-x>
- Nordenfelt, L. (2006). The concepts of health and illness revisited. *Medicine, Health Care and Philosophy*, 10(1), 5. <https://doi.org/10.1007/s11019-006-9017-3>
- Oliver, M. (1996). *Understanding disability: From theory to practice* (pp. x, 192). New York, NY, US: St Martin's Press. <https://doi.org/10.1007/978-1-349-24269-6>
- Powell, R., & Scarffe, E. (2019a). "Rethinking 'Disease': a fresh diagnosis and a new philosophical treatment." *Journal of Medical Ethics*, 45(9), 579–588. <https://doi.org/10.1136/medethics-2019-105465>
- Powell, R., & Scarffe, E. (2019b). Rehabilitating "Disease": Function, Value, and Objectivity in Medicine. *Philosophy of Science*, 86(5), 1168–1178. <https://doi.org/10.1086/705520>
- Putnam, H. (1975). Philosophy and Our Mental Life. In *Mind, Language, and Reality*. Cambridge University Press.
- Reznek, L. (1987). *The Nature of Disease*. Routledge & Kegan Paul.
- Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Research. Brain Research Reviews*, 18(3), 247–291. [https://doi.org/10.1016/0165-0173\(93\)90013-p](https://doi.org/10.1016/0165-0173(93)90013-p)
- Rombach, M. P., Porter, M. A., Fowler, J. H., & Mucha, P. J. (2014). Core-Periphery Structure in Networks. *SIAM Journal on Applied Mathematics*, 74(1), 167–190. <https://doi.org/10.1137/120881683>
- Sadegh-Zadeh, K. (2001). The Fuzzy Revolution: Goodbye to the Aristotelian Weltanschauung. *Artificial Intelligence in Medicine*, 21(1–3), 1–25. [https://doi.org/10.1016/S0933-3657\(00\)00071-3](https://doi.org/10.1016/S0933-3657(00)00071-3)
- Sadegh-Zadeh, K. (2008). The Prototype Resemblance Theory of Disease. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, 33(2), 106–139. <https://doi.org/10.1093/jmp/jhn004>
- Sadowsky, J. (2021). Before and After Prozac: Psychiatry as Medicine, and the Historiography of Depression. *Culture, Medicine, and Psychiatry*, 45(3), 479–502. <https://doi.org/10.1007/s11013-021-09729-2>
- Sartorius, N., & World Health Organization (Eds.). (1983). *Depressive disorders in different cultures: report on the WHO collaborative study on standardized assessment of depressive disorders*. Geneva : [Albany, N.Y: World Health Organization ; available from WHO Publications Centre].
- Sass, L., Pienkos, E., Skodlar, B., Stanghellini, G., Fuchs, T., Parnas, J., & Jones, N. (2017). EAWE: Examination of Anomalous World Experience. *Psychopathology*, 50, 10–54. <https://doi.org/10.1159/000454928>
- Sayette, M. A. (2016). The Role of Craving in Substance Use Disorders: Theoretical and Methodological Issues. *Annual Review of Clinical Psychology*, 12, 407–433. <https://doi.org/10.1146/annurev-clinpsy-021815-093351>
- Schmittmann, V. D., Cramer, A. O. J., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, 31(1), 43–53. <https://doi.org/10.1016/j.newideapsych.2011.02.007>
- Serre, F., Fatseas, M., Denis, C., Swendsen, J., & Auriacombe, M. (2018). Predictors of craving and substance use among patients with alcohol, tobacco, cannabis or opiate addictions: Commonalities and specificities across substances. *Addictive Behaviors*, 83, 123–129. <https://doi.org/10.1016/j.addbeh.2018.01.041>

- Shah, J. L., Scott, J., McGorry, P. D., Cross, S. P. M., Keshavan, M. S., Nelson, B., et al. (2020). Transdiagnostic clinical staging in youth mental health: a first international consensus statement. *World Psychiatry, 19*(2), 233–242. <https://doi.org/10.1002/wps.20745>
- Shapiro, S., Skinner, E. A., Kessler, L. G., Von Korff, M., German, P. S., Tischler, G. L., et al. (1984). Utilization of health and mental health services. Three Epidemiologic Catchment Area sites. *Archives of General Psychiatry, 41*(10), 971–978. <https://doi.org/10.1001/archpsyc.1984.01790210053007>
- Spitzer, R. L. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry, 24*(5), 399–411. [https://doi.org/10.1016/0010-440X\(83\)90032-9](https://doi.org/10.1016/0010-440X(83)90032-9)
- Spitzer, R. L. (1999). Harmful dysfunction and the DSM definition of mental disorder. *Journal of Abnormal Psychology, 108*, 430–432. <https://doi.org/10.1037/0021-843X.108.3.430>
- Spitzer, & Wakefield. (1999). DSM-IV diagnostic criterion for clinical significance: does it help solve the false positives problem? *The American Journal of Psychiatry, 156*(12). <https://doi.org/10.1176/ajp.156.12.1856>
- Stein, D. J. (2021). *Problems of Living: Perspectives from Philosophy, Psychiatry, and Cognitive-Affective Science* (1st edition.). San Diego: Academic Press.
- Stelling, J., Sauer, U., Szallasi, Z., Doyle, F., & Doyle, J. (2004). Robustness of Cellular Functions. *Cell, 118*, 675–85. <https://doi.org/10.1016/j.cell.2004.09.008>
- Szasz, T. S. (1960). The myth of mental illness. *American Psychologist, 15*(2), 113–118. <https://doi.org/10.1037/h0046535>
- Tabb, K. (2019). Philosophy of psychiatry after diagnostic kinds. *Synthese, 196*. <https://doi.org/10.1007/s11229-017-1659-6>
- Tekin, Ş. (2016). Are Mental Disorders Natural Kinds?: A Plea for a New Approach to Intervention in Psychiatry. *Philosophy, Psychiatry, & Psychology, 23*(2), 147–163. <https://doi.org/10.1353/ppp.2016.0013>
- Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again* (Illustrated edition.). New York: Basic Books.
- Tsou. (2007). Hacking on the Looping Effects of Psychiatric Classifications: What Is an Interactive and Indifferent Kind? *International Studies in the Philosophy of Science, 21*(3), 329–344. <https://doi.org/10.1080/02698590701589601>
- Tsou. (2015). DSM-5 and Psychiatry's Second Revolution: Descriptive Vs. Theoretical Approaches to Psychiatric Classification. In Steeves Demazeux & P. Singy (Eds.), *The DSM-5 in Perspective: Philosophical Reflections on the Psychiatric Babel* (pp. 43–62). Springer.
- Tsou. (2016). Natural kinds, psychiatric classification and the history of the DSM. *History of Psychiatry, 27*. <https://doi.org/10.1177/0957154X16656580>
- Tsou, J. Y. (2021a). *Philosophy of Psychiatry*. S.l.: Cambridge University Press.
- Tsou, J. Y. (2021b). *Philosophy of Psychiatry*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108588485>
- Tsou, J. Y. (2022). Biological Essentialism, Projectable Human Kinds, and Psychiatric Classification. *Philosophy of Science, 1*–21. <https://doi.org/10.1017/psa.2022.64>
- Vafaie, N., & Kober, H. (2022). Association of Drug Cues and Craving With Drug Use and Relapse: A Systematic Review and Meta-analysis. *JAMA psychiatry, 79*(7), 641–650. <https://doi.org/10.1001/jamapsychiatry.2022.1240>
- van Borkulo, C., Boschloo, L., Borsboom, D., Penninx, B. W. J. H., Waldorp, L. J., & Schoevers, R. A. (2015). Association of Symptom Network Structure With the Course of [corrected] Depression. *JAMA psychiatry, 72*(12), 1219–1226. <https://doi.org/10.1001/jamapsychiatry.2015.2079>
- Wachbroit, R. (1994). Normality as a Biological Concept. *Philosophy of Science, 61*(4), 579–591.
- Wakefield. (1992). The concept of mental disorder. On the boundary between biological facts and social values. *The American Psychologist, 47*(3), 373–388. <https://doi.org/10.1037//0003-066x.47.3.373>
- Wakefield. (1999). Mental disorder as a black box essentialist concept. *Journal of Abnormal Psychology, 108*(3), 465–472. <https://doi.org/10.1037/0021-843X.108.3.465>
- Wakefield, J. C. (2006). Are there relational disorders? A harmful dysfunction perspective: Comment on the special section. *Journal of Family Psychology, 20*(3), 423–427. <https://doi.org/10.1037/0893-3200.20.3.423>
- Wakefield, J. C. (2012). Are you as smart as a 4th grader? Why the prototype-similarity approach to diagnosis

- is a step backward for a scientific psychiatry. *World Psychiatry*, 11(1), 27–28.
- Wakefield, J. C. (2013). The DSM-5 debate over the bereavement exclusion: Psychiatric diagnosis and the future of empirically supported treatment. *Clinical Psychology Review*, 33(7), 825–845. <https://doi.org/10.1016/j.cpr.2013.03.007>
- Wakefield, J. C., & Conrad, J. A. (2019). Does the harm component of the harmful dysfunction analysis need rethinking?: Reply to Powell and Scarffe. *Journal of Medical Ethics*, 45(9), 594–596. <https://doi.org/10.1136/medethics-2019-105578>
- Wakefield, J. C., & First, M. B. (2013). Diagnostic Validity and the Definition of Mental Disorder: A Program for Conceptually Advancing Psychiatry. *The Canadian Journal of Psychiatry*, 58(12), 653–655. <https://doi.org/10.1177/070674371305801201>
- Wakefield, J. C., & Spitzer, R. L. (2002). *Why requiring clinical significance does not solve epidemiology's and DSM's validity problem: Response to Regier and Narrow* (p. 40). Arlington, VA, US: American Psychiatric Publishing, Inc.
- Westen, D. (2012). Prototype diagnosis of psychiatric syndromes. *World Psychiatry*, 11(1), 16–21. <https://doi.org/10.1016/j.wpsyc.2012.01.004>
- World Health Organization. (2017). International Classification of Diseases 11th Edition (ICD-11). In *Encyclopedia of Personality and Individual Differences* (pp. 1–6). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-28099-8_904-1
- Zachar, P. (2000). Psychiatric Disorders Are Not Natural Kinds. *Philosophy, Psychiatry, and Psychology*, 7(3), 167–182.
- Zachar, P. (2014). *A Metaphysics of Psychopathology*. MIT Press.
- Zachar, P. (2015). Psychiatric disorders: natural kinds made by the world or practical kinds made by us? *World Psychiatry*, 14(3), 288–290. <https://doi.org/10.1002/wps.20240>
- Zachar, P. (2021). Psychiatric Disorders and the Imperfect Community: A Nominalist HDA. In *Defining Mental Disorder Jerome Wakefield and His Critics*. <https://direct.mit.edu/books/book/5015/chapter/2812055/Psychiatric-Disorders-and-the-Imperfect-Community>. Accessed 21 July 2022

Table des matières

Remerciements	4
Sommaire.....	6
Avertissements.....	7
Table des figures.....	9
Tableaux	9
Liste des abréviations	10
INTRODUCTION GÉNÉRALE	12
L'approche par la catégorie et la dimension.....	17
L'approche par le symptôme : présentation et intérêts.....	20
Thèse principale.....	23
Plan	31
LES MULTIPLES FONCTIONS DU MANUEL DIAGNOSTIQUE ET STATISTIQUE DES TROUBLES MENTAUX : UN REFLET DE SON HISTOIRE	35
Introduction	36
I. En amont du DSM.....	38
II. Les critères de Feighner et les critères diagnostiques pour la recherche	41
1. La CIM et le DSM.....	41
2. Les raisons de l'acceptation d'un nouveau système de classification pour la psychiatrie	44
3. L'élaboration des critères de Feighner	46
4. L'élaboration des critères diagnostiques pour la recherche	51
III. Les enjeux du DSM-III : entre opérationnalisation et contingence.....	53
1. Importance de l'opérationnalisation.....	53
2. Origine de l'opérationnalisation.....	56
3. Contextualisation socio-historique du DSM-III.....	57
IV. L'évolution des versions du DSM après le DSM-III	60
1. L'élaboration du DSM-IV	61
2. L'élaboration du DSM-5 : un processus conservateur	62
3. Implications de l'ancrage institutionnel du DSM en psychiatrie	65
Conclusion	68
LE PROJET <i>RESEARCH DOMAIN CRITERIA</i> : DES NEUROSCIENCES À L'ENVIRONNEMENT	72
Introduction	73
I. Description de la matrice des RDoC	79
II. Genèse des RDoC	84
1. L'influence du contexte scientifique de la première moitié du XX ^e siècle	84
2. Le contexte scientifique de la seconde moitié du XX ^e siècle : une triple déception.....	89

3.	L'influence de la technologie et des marqueurs biologiques	93
4.	Vers un nouveau système.....	98
III.	Validité de construit des RDoC	105
1.	Validité	106
2.	Validité de construit	108
IV.	Les RDoC et les trois niveaux d'intégration	112
1.	Intégrations mutuelles entre les principales fonctions cérébrales (entre les lignes de la matrice).....	114
2.	Intégrations mutuelles entre unités d'analyse (entre les colonnes de la matrice)	116
3.	Intégration sur le plan interdisciplinaire.....	119
4.	Conclusion sur l'intégration	120
V.	RDoC, intégration du neurodéveloppement et application clinique.....	121
1.	RDoC et comportements	122
2.	RDoC et neurodéveloppement	124
3.	RDoC et clinique.....	127
	Conclusion	130
	LA TAXONOMIE HIÉRARCHIQUE DE LA PSYCHOPATHOLOGIE : À LA CROISÉE ENTRE CLINIQUE ET STATISTIQUES	133
	Introduction	134
I.	Psychologie dimensionnelle, nosologie quantitative et statistiques	137
1.	La psychologie dimensionnelle.....	137
2.	La nosologie quantitative	139
3.	Classification numérique statistique.....	140
II.	La structure dimensionnelle de la HiTOP.....	142
1.	Description de la dimensionnalité.....	142
III.	La structure hiérarchique de la HiTOP	145
1.	Description de la hiérarchisation.....	145
2.	Hiérarchisation, clinique et comorbidité	147
IV.	Discussion des enjeux	149
1.	Enjeux de validité.....	149
2.	Enjeux méthodologiques	152
3.	Enjeux inhérents à la dimensionnalité.....	154
4.	Enjeux d'utilité clinique	155
	Conclusion	157
	DE L'APPROCHE PAR LA CATÉGORIE À L'APPROCHE PAR LE SYMPTÔME	169
	Introduction	170
	Présupposé I : Variables latentes et mécanismes locaux	172
1.	Approche par le symptôme	173
a.	Influence causale mutuelle	173
b.	Abandon de la variable latente	174

2.	Notion de noyau central	178
a.	Noyau central.....	178
b.	Composants du noyau central.....	181
3.	Notion de mécanisme	183
a.	Activations des mécanismes biologiques	183
b.	Mécanismes à l'œuvre au sein des réseaux	185
4.	Résumé	188
	Présupposé II : Préjudice et analyse du symptôme préjudiciable.....	189
1.	Rôle du préjudice en philosophie de la psychiatrie.....	189
2.	Notion de symptôme préjudiciable	193
3.	Préjudices et espèce psychiatrique	196
4.	Réseau figé.....	199
	Une définition potentiellement pertinente de ce qu'est une espèce psychiatrique pour la pratique clinique.....	202
1.	Avantage de considérer des espèces psychiatriques utiles pour le clinicien.....	202
2.	Proposition de définition d'une espèce psychiatrique.....	205
	Conclusion.....	209
	DES CATÉGORIES AUX SYMPTÔMES.....	215
1.	Implications pour le clinicien et la nosologie.....	224
2.	Quelle sorte d'espèce psychiatrique constitue ce noyau central ?.....	225
3.	Implications du noyau central du réseau pour la stabilité de la nosologie	226
4.	Noyau central en philosophie de la médecine	226
	Bibliographie	233
	DSM.....	233
	RDoC.....	237
	HiTOP.....	243
	Approche par le symptôme.....	247
	Table des matières	255
	Classifications psychiatriques : des catégories aux symptômes	258
	Résumé	258
	Mots-clés	258
	Summary.....	259
	Keywords.....	259

Classifications psychiatriques : des catégories aux symptômes

Résumé

Ces dernières décennies, la philosophie de la psychiatrie s'est principalement appuyée sur les catégories et dimensions présentées au sein de classifications internationales comme le manuel diagnostique et statistique des troubles mentaux (*Diagnostic and Statistical Manual of mental disorders* – DSM). Une telle classification, à l'image d'autres conceptions émergentes comme le projet *Research Domain Criteria* (RDoC) ou la Taxonomie Hiérarchique de la Psychopathologie (*Hierarchical Taxonomy of Psychopathology* – HiTOP) intègre à la définition d'une espèce psychiatrique l'hypothèse d'une variable latente (une entité sous-jacente au phénotype, non directement observable). Or, une telle variable latente se révèle peu pertinente pour guider le clinicien dans sa pratique clinique (notamment diagnostique et thérapeutique).

Nous admettons que pour satisfaire les multiples objectifs auxquels doit répondre la définition d'une espèce psychiatrique (par exemple, administratifs, financiers ou politiques), une définition fondée sur la notion de variable latente peut être pertinente. L'analyse de la dysfonction préjudiciable semble être un bon candidat de définition utile à ces objectifs non-cliniques.

Cependant, notre thèse est que cette approche fondée sur la notion de variable latente échoue à répondre aux objectifs de la pratique clinique. Nous proposons une définition alternative de ce qu'est une espèce psychiatrique, fondée sur les symptômes. Dans le cadre de cette définition, les symptômes, qui rendent compte cliniquement des préjudices vécus par les patients, constituent les produits de mécanismes activés. L'ensemble organisé de ces symptômes en interactions mutuelles permet de fournir à la pratique clinique une espèce psychiatrique pertinente.

Mots-clés

Psychiatrie ; classification ; philosophie ; symptôme ; espèces psychiatriques ; nosologie ; diagnostic ; épistémologie.

Summary

In recent decades, philosophy of psychiatry has mainly relied on categories and dimensions presented in international classifications such as the Diagnostic and Statistical Manual of Mental Disorders (DSM). Such a classification, like other emerging taxonomies such as the Research Domain Criteria (RDoC) project or the Hierarchical Taxonomy of Psychopathology (HiTOP) are based on a category-based approach, which integrates into the definition of a psychiatric kind the hypothesis of a latent variable (i.e., a not directly observable entity).

Our main thesis is to show that this multiplicity of conceptions of psychiatry would be related to the desire to respond to both clinical and non-clinical functions (e.g., institutional, political or financial) – a thesis having led these different conceptions of psychiatry to adopt an approach by category. This approach, based on the notion of latent variable, seems not really relevant to the clinician in his/her clinical practice.

The definition of a psychiatric kind help to respond to various objectives. First, a definition based on the notion of latent variable may prove to be relevant for most of these purposes. More precisely, the harmful dysfunction analysis seems to be a robust definition candidate useful for non-clinical purposes.

However, this approach based on the notion of latent variable fails to meet the objectives of clinical practice. In the particular situation of clinical practice, we propose another definition of a psychiatric kind, which would be based directly on symptoms. We thus propose a definition in which the symptoms, reflecting patient harms, are the products of activated mechanisms in mutual interactions. Such an organized symptom set accounts for a psychiatric kind relevant to clinical practice.

Assuming the adoption of at least two definitions is necessary to meet the clinical and non-clinical challenges of psychiatry. This opens the way to a reconciliation between different conceptions of psychiatry, through a tolerant respect for a multiplicity of definitions of a psychiatric disorder.

Keywords

Psychiatry; classification; philosophy; symptom; psychiatric kind; nosology; diagnostic; epistemology.

Discipline : Philosophie

École Doctorale : École Doctorale de Philosophie, Université Paris 1 Panthéon-Sorbonne, 13, rue du Four, 75006 Paris.

Équipes d'accueil (codirection) : Institut d'Histoire et de Philosophie des Sciences et des techniques (UMR8590), 13 rue du Four, 75006 Paris / Institut de Recherches Philosophiques de Lyon, 18 rue Chevreul, 69007 Lyon

