

Pourquoi faut-il réviser le critère explicatif de la HDA?

MAXIME F. GIGUÈRE, *Université Laval*

RÉSUMÉ : Ce texte offre une lecture critique du critère explicatif de la HDA (*harmful dysfunction analysis*) proposée par Jerome C. Wakefield¹. À l'origine, la HDA avait pour ambition d'offrir une issue au débat entre les positions normativistes et naturalistes en proposant une analyse hybride du concept de trouble mental. Or comme nous le montrons ici, le critère explicatif de la HDA rencontre deux difficultés issues de l'adoption par Wakefield d'une perspective évolutionniste. La première difficulté relève de l'impossibilité à définir le critère explicatif en termes strictement factuels, ce qui rend le critère en partie évaluatif et de manière globale fait pencher la HDA du côté normativiste. La seconde difficulté quant à elle relève de l'impossibilité qu'il y a à déterminer quels états pathologiques remplissent les conditions du critère en question. En définitive, nous défendons la thèse selon laquelle la solution à privilégier pour sauver la HDA consiste à réviser le critère explicatif, notamment en écartant la dimension évolutionniste sur laquelle il repose.

Introduction

Les premières tentatives pour définir le concept de trouble mental devaient montrer que ce concept était significatif et que, contrairement à ce qui était affirmé par l'antipsychiatrie, la maladie mentale n'était pas un mythe. Une telle impulsion est particulièrement manifeste dans les textes de Boorse² et de Kendell³.

Le projet définitionnel a par la suite été repris par des philosophes de la psychiatrie cherchant à formuler une analyse conceptuelle satisfaisante du trouble mental. L'analyse conceptuelle est l'une des méthodes employées en philosophie des sciences ; elle consiste en l'articulation des concepts clefs d'une théorie scientifique⁴. Cette méthode permettrait de déterminer les valeurs épistémiques, les limites, les biais et le pouvoir explicatif d'une théorie scientifique⁵. L'analyse conceptuelle du trouble mental la plus influente aujourd'hui est la *Harmful dysfunction analysis* (HDA) proposée par le philosophe Jerome Wakefield⁶. Selon cette analyse, «trouble mental» est un concept hybride qui comprend un critère explicatif (*dysfunction*) et un critère évaluatif (*harmful*).

Selon nous, une telle analyse demeure insatisfaisante puisque sa dimension factuelle, c'est-à-dire son critère explicatif, fait intervenir des valeurs. Cette difficulté proviendrait de la thèse évolutionniste, dont la conséquence serait l'emploi d'un concept évaluatif de dysfonction. La première section de cet article présentera le contexte intellectuel dans lequel la HDA a été proposée afin de comprendre les difficultés auxquelles devait pallier une approche hybride. La section suivante présentera plus en détail les deux critères de la HDA en insistant sur la thèse évolutionniste qui sous-tend le critère explicatif. La dernière section, enfin, cherchera à montrer qu'en vertu de cette thèse, le concept de dysfonction - central pour le critère explicatif - n'est pas rigoureusement factuel et ne permet pas de distinguer le trouble de l'état normal. En conclusion, nous pensons que ces deux difficultés de la HDA justifient le recours à un concept non-évolutionniste de dysfonction dans la révision de cette analyse conceptuelle.

1. Le contexte d'apparition de la HDA

Présenter le contexte d'apparition de la HDA permet de comprendre les intentions qui l'ont fait naître. Cela offre aussi l'occasion d'illustrer les raisons qui justifient l'emploi d'une analyse conceptuelle qui ne soit ni naturaliste, ni normativiste. Au début des années 90, lorsque Wakefield formule la HDA, le débat

entre les positions naturalistes et normativistes du trouble mental domine la littérature en philosophie de la médecine⁷. Ces positions se présentent comme étant mutuellement exclusives : pour elles, le « trouble mental » est ou bien un concept factuel, ou bien un concept évaluatif. Aucune commune mesure ne semble alors possible entre ces positions. L'impossibilité de trouver une définition consensuelle du trouble mental et la multiplication d'exemples problématiques comme l'illettrisme ou l'anomalie bénigne aurait précipité la formulation d'autres types d'analyses conceptuelles dans les années 90, dont la HDA.

Selon l'analyse naturaliste, le concept de « trouble mental » réfère à un état pathologique (*condition*) factuellement et naturellement distinct de l'état normal. Cette analyse s'appuie généralement sur l'approche évolutionniste, pour laquelle le trouble mental est une dysfonction biologique. Le trouble mental serait causé par des mécanismes mentaux dysfonctionnels, c'est-à-dire dont l'opération ne serait pas conforme à la raison pour laquelle ils auraient été naturellement sélectionnés dans le passé⁸. L'analyse conceptuelle de Boorse constitue l'exemple le plus influent d'approche naturaliste⁹, même si son approche n'est pas évolutionniste. Boorse croit qu'il faut distinguer le concept pratique, la maladie (*illness*) du concept théorique, la pathologie (*disease*)¹⁰. La santé, notion à laquelle on oppose le concept théorique de pathologie, serait un fonctionnement normal où « la normalité est statistique et les fonctions sont biologiques¹¹ ». Une telle analyse conceptuelle ne laisserait aucune place aux valeurs des chercheurs et chercheuses et permettrait un concept strictement factuel de trouble mental en s'appuyant sur les données scientifiques. Les valeurs n'auraient alors aucune influence sur la détermination des états pathologiques.

Pour ce qui est de l'analyse normativiste, elle suggère plutôt que le concept de trouble mental relève d'un jugement de valeur. Le terme de « trouble mental » référerait à un état jugé indésirable et viendrait affirmer qu'un état est qualifié de pathologique lorsqu'il est généralement évalué négativement. L'analyse conceptuelle de Sedgwick selon laquelle « la maladie est essentiellement une

déviations [*deviancy*] à partir d'un état jugé plus désirable¹²» est un exemple d'analyse normativiste. Qualifier une condition de trouble mental selon cette analyse repose sur une comparaison entre la désirabilité de deux états. Une analyse normativiste ne laisse aucune place à la présence de faits déterminants, puisqu'ultimement ce qui est une maladie dépend de l'évaluation individuelle. La stérilité constitue un exemple paradigmatique de ce type d'analyse, puisque des interventions médicales comme la vasectomie ou la fécondation in vitro montrent qu'un même état peut être désiré ou évité, illustrant ainsi le caractère évaluatif du concept de maladie.

Malgré leur opposition, ces deux analyses conceptuelles cherchent à rendre compte du concept de «trouble mental» par un seul critère. S'excluant mutuellement, cette opposition conduit à l'impossibilité de parvenir à un consensus qui permettrait de résoudre les enjeux pratiques que devait clarifier l'analyse conceptuelle¹³. Autrement dit, de l'opposition frontale entre deux analyses mutuellement exclusives serait née, en raison de l'impossibilité à atteindre un consensus, une «controverse stérile¹⁴» menant ainsi à une insatisfaction générale face à la méthode de l'analyse conceptuelle. Un ensemble de cas problématiques touchant chacune des analyses contribuera aussi à cette insatisfaction généralisée. Les analyses normativistes font face à la difficulté de distinguer les troubles mentaux d'autres conditions évaluées négativement telles que l'ignorance, le manque d'adresse, le manque de talent, une faible intelligence, l'illettrisme, la criminalité, les mauvaises manières, la bêtise et la faiblesse morale¹⁵. En effet, une analyse conceptuelle qui définit le trouble comme une déviation à partir d'un état désirable explique difficilement pourquoi des états comme ceux que nous venons de nommer ne sont pas des troubles alors qu'ils sont jugés indésirables. L'approche naturaliste peine quant à elle à expliquer l'anomalie bénigne. Une déviation statistique qui ne se traduit par aucune conséquence préjudiciable chez l'individu devra être jugée pathologique, alors qu'une telle conclusion ne semble pas être l'intuition professionnelle¹⁶. L'impossibilité d'atteindre un consensus et l'accumulation d'exemples problématiques mèneront à une insatisfaction face à ces deux approches.

Cette insatisfaction précipitera la multiplication d'approches désireuses de se constituer en une troisième voie. Ces nouvelles analyses conceptuelles auront comme prétention d'éviter les exemples problématiques présentés ci-haut et de permettre un consensus. Une analyse proposée par Lilienfeld et Marino analyse le « trouble mental » comme un concept prototypique¹⁷. Selon eux, le « trouble mental » est un concept dont les frontières ne peuvent pas être clairement définies avec des critères nécessaires et suffisants¹⁸. Cette analyse s'inspire des jeux de langage wittgensteinien où il n'existe aucune règle nécessaire et suffisante pour qu'une activité soit considérée comme un « jeu ». Une autre analyse, développée par Reznek, définit le trouble mental par la conjonction de plusieurs critères nécessaires¹⁹. Selon Reznek, un trouble serait un état qui nécessite une intervention médicale et qui est préjudiciable pour un membre standard d'une communauté²⁰. Une telle analyse conceptuelle fait appel à deux critères nécessaires plutôt qu'à un seul pour définir le trouble mental. Enfin, la HDA développée par Wakefield met en rapport le critère explicatif de dysfonction avec le critère évaluatif de préjudice.

2. Les deux critères de la HDA (Harmful dysfunction analysis)

La HDA est d'abord développée dans deux textes publiés en 1992. Elle est une analyse conceptuelle hybride selon laquelle la conjonction d'un critère explicatif et d'un critère évaluatif permettrait de rendre compte du « trouble mental ». La HDA est le plus explicitement présentée dans ce passage :

Un état est un trouble pathologique si et seulement si (a) cet état produit un certain préjudice ou prive la personne d'un bénéfice au regard des attentes de la culture à laquelle elle appartient (critère évaluatif), et (b) cet état résulte d'une incapacité d'un mécanisme interne à réaliser sa fonction naturelle, où par fonction naturelle nous entendons un effet qui fait partie de l'explication évolutionniste de l'existence et de la structure du mécanisme en question (critère explicatif)²¹.

La structure hybride de la HDA est manifeste dans ce passage : la conjonction d'un critère évaluatif et d'un critère explicatif serait nécessaire et suffisante à la définition du trouble mental. Le premier critère est évaluatif puisqu'il dépend de jugements à propos des faits, alors que le second est explicatif puisqu'il dépend des faits eux-mêmes. Wakefield s'appuie sur la distinction classique entre les faits et les valeurs et propose un critère à chacun de ces niveaux²².

2.1. Le critère évaluatif (a) : l'état doit porter préjudice

Le premier critère est dit évaluatif puisqu'il dépend d'une interprétation plutôt que seulement des faits. Pour le critère évaluatif de la HDA (*Harmful dysfunction*), c'est le préjudice (*harm*) qui constitue la condition nécessaire pour qu'un état soit jugé comme un trouble. «Préjudice» est un terme évaluatif qui renvoie aux conséquences subies et jugées négatives selon les standards de la collectivité²³. Ce sont les préjudices subis par l'individu plutôt que par la communauté qui sont significatifs pour ce critère. Les conséquences subies sont l'ensemble des symptômes qui découlent directement d'un état : la fatigue provoquée par la grippe constitue une conséquence subie, par exemple. La conséquence doit provenir directement de l'état pour qu'elle constitue un préjudice au sens de la HDA. Ainsi, la stigmatisation subie par une personne ne s'avère pas être un préjudice, puisqu'elle provient de la société plutôt que de l'état.

Pour Wakefield, l'évaluation des conséquences subies dépend des «standards de la collectivité»²⁴. La conséquence subie doit être généralement considérée comme indésirable au sein d'une culture pour constituer un préjudice. Ce sont ainsi les valeurs sociales plutôt qu'individuelles qui permettent de déterminer s'il y a préjudice. Examinons l'exemple de l'infertilité : même si un individu peut se réjouir d'un tel état, celui-ci est généralement jugé préjudiciable selon les standards de la culture. Cet état constitue alors un préjudice puisqu'il est évalué comme tel par la collectivité et indépendamment de l'évaluation individuelle. Tout état préjudiciable dans ce sens remplit le critère évaluatif de la HDA.

2.2. Le critère explicatif (b) : l'état doit provenir d'une dysfonction

Le critère explicatif stipule que l'état doit provenir d'une dysfonction d'un mécanisme interne, soit d'une déviation par rapport à la fonction naturelle, pour être pathologique. La conception de la fonction employée est dite étiologique et sera développée par Neander en s'appuyant sur les travaux de Larry Wright²⁵. La fonction naturelle correspondrait à l'effet qui explique la présence du mécanisme. Elle serait le résultat du processus d'adaptation par sélection naturelle. L'usage de la fonction naturelle est limité aux mécanismes et exclut une utilisation dérivée, par exemple pour qualifier une relation de dysfonctionnelle²⁶. L'impossibilité pour un mécanisme d'accomplir sa fonction naturelle remplit le critère explicatif de la HDA.

L'attribution d'une fonction naturelle dépend de sa capacité à expliquer la présence et la structure d'un mécanisme interne²⁷. Par exemple, l'évitement du danger serait la fonction attribuée au mécanisme de peur. Puisque l'évitement du danger explique la contribution du mécanisme de peur au succès évolutif, elle constitue la meilleure explication de la présence d'un tel mécanisme chez l'humain. Pour cette raison, l'évitement du danger est une fonction attribuée au mécanisme de peur. Pour illustrer le rôle explicatif de la fonction, Wakefield emploie l'analogie de l'artefact²⁸. Chaque partie de l'artefact s'explique par sa fonction au sein du tout : la présence d'une roue sur un véhicule s'explique par sa contribution au succès du tout. Cette fonction fournit la meilleure explication de la présence de la roue. Cette analogie n'est toutefois pas entière, puisque seul l'artefact a un concepteur explicite²⁹, alors que la conception des organismes serait le résultat de la sélection naturelle.

Wakefield défend une thèse évolutionniste selon laquelle la sélection naturelle fournit la meilleure explication de la présence de mécanismes mentaux et de leur fonction³⁰. Selon lui, la présence de tels mécanismes ne pourrait être due au hasard ; la sélection naturelle est « une découverte scientifique capitale, et non une question de définition³¹ ». Ces mécanismes favoriseraient la survie et la reproduction, qui seraient les buts du vivant indépendants

de toute valeur, constituant ainsi le critère explicatif de la HDA. Wakefield le résume de cette manière :

Pour le dire brièvement, les mécanismes qui se sont trouvés, il y a longtemps, avoir favorisé par leurs effets le succès reproductif de certains organismes sur un nombre suffisamment de générations ont vu leur fréquence augmenter; ils ont été «naturellement sélectionnés» et c'est ce qui explique leur existence dans les organismes présents. Par conséquent, expliquer un mécanisme par sa fonction naturelle est en quelque sorte la même chose, mais à l'envers, que de l'expliquer de manière causale par le processus de sélection naturelle³².

Les mécanismes mentaux présents chez l'humain auraient alors été sélectionnés par leur contribution au succès reproductif.

Une telle position suppose l'existence de mécanismes mentaux et d'un certain *design* de l'esprit. Ce *design* serait découvert par une méthode, à savoir, l'analyse fonctionnelle évolutive, et consisterait à «inférer les causes proximales du comportement (le fonctionnement des mécanismes proximaux) à partir de prémisses portant sur les causes ultimes de notre comportement³³». Autrement dit, la méthode consisterait à supposer des avantages aux mécanismes mentaux passés afin de connaître leur fonction actuelle. Ces fonctions demeureraient toutefois toujours hypothétiques : pour des raisons épistémologiques, il ne sera jamais possible par principe d'effectuer l'inventaire des mécanismes mentaux. En effet, certains mécanismes spécifiques à l'espèce et produits lors du Pléistocène ne pourront jamais être découverts empiriquement³⁴. L'on ne peut donc que supposer qu'il existe un éventail de mécanismes mentaux chez l'humain et que celui-ci aurait été façonné de manière relativement uniforme par l'évolution³⁵. De telles affirmations ne peuvent pas, par principe, dépasser un statut hypothétique. Une telle thèse évolutionniste est défendue par Wakefield et sous-tend aussi la majorité des analyses conceptuelles naturalistes³⁶.

3. Deux problèmes avec la thèse évolutionniste

Nous avançons que la thèse évolutionniste ne permet pas de fournir un critère explicatif satisfaisant pour le projet explicatif de la HDA. Nous considérons, comme plusieurs auteurs³⁷, que la thèse évolutionniste est la principale difficulté de la HDA. D'une part, le concept de dysfonction serait défini en termes évaluatifs. Pour cette raison, le critère explicatif serait en fait évaluatif et la HDA ne serait pas véritablement hybride. D'autre part, il serait difficile d'identifier les mécanismes mentaux et leur fonction naturelle, puisqu'il existe plusieurs hypothèses évolutionnistes probables. Sans une connaissance précise des mécanismes mentaux, la présence de dysfonction serait toujours indéterminée, rendant alors le critère explicatif inutilisable.

3.1. « Dysfonction » est un concept évaluatif

La première difficulté de la thèse évolutionniste est que le concept central de dysfonction serait évaluatif. Il ne constituerait pas un critère seulement factuel, rendant ainsi la HDA normativiste plutôt qu'hybride. À titre indicatif, il peut être utile de distinguer le projet descriptif du projet explicatif de Wakefield. Le projet descriptif vise à « rendre compte des jugements classificatoires largement partagés³⁸ ». Or, le concept de fonction naturelle, dans une perspective évolutionniste, est généralement considéré évaluatif³⁹. Le concept de dysfonction employé dans un sens factuel par Wakefield ne permettrait alors pas d'atteindre les objectifs du projet descriptif. Sur le terrain du projet descriptif, Wakefield serait forcé d'admettre que « dysfonction » est un concept évaluatif et que la HDA n'est pas véritablement hybride. Nous croyons donc qu'il faille plutôt s'intéresser au projet explicatif de Wakefield qui vise à formuler un énoncé substantiel sur ce qui constitue véritablement un trouble⁴⁰. Toutefois, nous croyons qu'une difficulté est aussi présente sur le projet explicatif : l'impossibilité de définir la dysfonction en termes factuels indiquerait que le concept de dysfonction est évaluatif.

Selon une exigence couramment employée dans la littérature, un terme doit être défini en termes seulement factuels pour être jugé factuel⁴¹. Autrement, le concept doit être jugé évaluatif. Comme le critère explicatif de la HDA se veut factuel, la dysfonction doit pouvoir être définie en termes seulement factuels. Wakefield juge qu'une telle exigence est pertinente pour déterminer la factualité d'un concept puisqu'il l'emploie lui-même pour critiquer le *Manuel diagnostique et statistique des troubles mentaux* (DSM)⁴², qui est l'ouvrage de référence de la psychiatrie. Wakefield affirme que le concept central de « trouble mental » du *DSM-III-R* (troisième version révisée du manuel) est évaluatif plutôt que factuel puisqu'il n'est pas formulé en termes factuels⁴³. L'impossibilité d'une définition en termes seulement factuels en ferait un concept évaluatif. L'emploi d'une critique par Wakefield illustre son admission qu'une telle exigence est pertinente pour évaluer la factualité d'un concept. Nécessairement, Wakefield devrait donc pouvoir définir seulement en termes factuels le concept de dysfonction.

Or même si Wakefield croit que le concept de dysfonction est factuel, il ne réussit pas à remplir cette exigence. Ni le concept de dysfonction, ni le concept opposé, à savoir celui de fonction naturelle, ne sont définis en termes uniquement ou strictement factuels⁴⁴. Dans un texte fondateur, Wakefield emploie un vocabulaire évaluatif pour décrire les fonctions des mécanismes naturels : « Les bénéfiques que procurent les mécanismes naturels, comme les artefacts, sont si remarquables et reposent sur des interactions *si harmonieuses et si intriquées* qu'il est souvent raisonnable de considérer que le bénéfice n'est pas accidentel⁴⁵ ». Selon l'exigence définitionnelle, les fonctions naturelles de ces mécanismes ne peuvent pas être factuelles, puisqu'elles sont définies en termes évaluatifs. En effet, l'« harmonie » étant un terme évaluatif, les fonctions des mécanismes seraient nécessairement évaluatives. Le critère explicatif de la HDA s'appuierait donc sur un concept évaluatif de « fonction naturelle » et constituerait plutôt un second critère évaluatif.

Afin d'éviter une telle difficulté (mise en évidence par Sadler et Agich⁴⁶), Wakefield fournit une reformulation de son propos :

A natural function of a biological mechanism is an effect of the mechanism that is part of the explanation of the existence or structure of the mechanism; and a biological mechanism is dysfunctional when it does not perform one of its natural function⁴⁷.

Malgré cette reformulation, nous croyons que la fonction naturelle demeure un concept évaluatif. L'«explication» sur laquelle s'appuierait une telle reformulation ne serait pas seulement factuelle puisqu'elle dépendrait d'au moins trois choix théoriques. En raison de ces choix théoriques, l'«explication» ne serait pas un terme factuel et la dysfonction ne serait alors pas définie en termes seulement factuels. L'exigence définitionnelle échouerait malgré la reformulation.

Un premier choix théorique opéré par Wakefield est l'adoption d'une explication évolutionniste des fonctions naturelles. Pour Wakefield, «la théorie évolutionniste fournit la seule explication scientifique plausible qui existe actuellement de la façon dont les fonctions normales d'un mécanisme peuvent expliquer l'existence et la structure de celui-ci⁴⁸». Contrairement à ce qu'affirme Wakefield, la théorie de l'évolution ne serait toutefois pas la seule explication possible des mécanismes mentaux. En effet, l'explication fonctionnelle pourrait aussi s'appuyer sur un critère biochimique, social ou anatomique⁴⁹. La théorie de l'évolution n'est pas la seule théorie scientifique à employer le concept de dysfonction. Le concept de dysfonction tirée de la théorie de l'évolution ne serait donc pas le seul permettant de fournir une explication satisfaisante de la présence de mécanismes mentaux. Un physiologiste, par exemple, fournirait une explication fonctionnelle des mécanismes mentaux bien différente de celle choisie par Wakefield. Même si de bonnes justifications peuvent être présentées pour défendre une conception évolutionniste de la dysfonction⁵⁰, il n'en demeure pas moins que l'adoption d'une telle conception est un choix, et qu'ainsi

le fait de favoriser une telle explication est un choix théorique effectué par Wakefield.

Toutefois, quand bien même la théorie de l'évolution serait la seule explication fonctionnelle possible, un second choix théorique s'opère au sein même de la théorie de l'évolution, puisque la survie et la reproduction ont été choisies comme buts théoriques par les biologistes de l'évolution. Comme Gammelgaard l'a signalé, la survie et la reproduction sont considérées les buts du vivant au sein de la théorie de l'évolution parce que ce sont les phénomènes que cherchent à expliquer les chercheurs⁵¹. Autrement dit, il y a déjà, au sein même de la théorie de l'évolution, un choix théorique⁵². Même si l'explication évolutionniste était la seule explication fonctionnelle, elle serait tout de même le résultat de choix théoriques.

Enfin, un troisième choix théorique consiste à identifier les buts des organismes aux buts de la théorie de l'évolution. Expliquer le trouble médical par la « dysfonction » et employer en médecine la théorie de l'évolution est un troisième choix théorique. Or, une telle équivalence n'est toutefois pas nécessaire. En effet, la biologie s'intéresse à plusieurs activités de l'organisme qui ne sont pas axées vers la survie et la reproduction, par exemple la sexualité non reproductive⁵³. Ainsi, la théorie de l'évolution n'indiquerait pas nécessairement que les troubles doivent être définis en termes évolutifs. En effet, les activités humaines n'étant pas toutes axées vers la survie et la reproduction, expliquer les troubles médicaux par la dysfonction serait un choix théorique⁵⁴. Même si Wakefield présente de bonnes raisons d'employer un tel cadre, une définition de la santé par la biologie évolutionniste demeure un choix théorique.

En somme, l'explication favorisée par Wakefield dépend de choix théorique plutôt que de nécessité pratique. En définissant la dysfonction par l'explication de la présence de mécanismes, Wakefield ne formule pas une définition véritablement exempte de termes évaluatifs. L'« explication » serait un terme évaluatif puisque l'explication favorisée par Wakefield dépend de choix auxquels sont imbriquées des valeurs. Nous avons d'abord présenté l'exigence définitionnelle acceptée par Wakefield et l'échec d'une première

définition de la dysfonction. L'échec de cette seconde tentative confirme que «dysfonction» est un terme évaluatif. Il en découle que le critère de dysfonction n'est pas un seulement factuel : la HDA n'est donc pas véritablement hybride. S'appuyant ainsi sur des critères évaluatifs, elle serait sujette aux mêmes difficultés que les analyses normativistes du trouble mental.

3.2. L'impossibilité d'identifier les mécanismes mentaux

Le second problème de la thèse évolutionniste est la difficulté à identifier quels sont les mécanismes mentaux et quelles sont leurs fonctions naturelles. Sans connaissance des mécanismes mentaux, il demeure impossible de déterminer à quel moment il y a dysfonction. Le critère explicatif serait ainsi inutilisable, ne permettant pas déterminer quels états sont pathologiques. Cette difficulté proviendrait de la coexistence de trois explications évolutionnistes plausibles et incompatibles : le *mismatch*, le *breakdown* et la *persistence*. Toutes trois expliquent la présence des mécanismes à partir de leur fonction naturelle⁵⁵. Or, il n'est pas toujours clair laquelle fournit la meilleure explication de la présence des mécanismes mentaux. Notre présentation des trois hypothèses montrera qu'elles ont une capacité explicative équivalente, notamment pour expliquer le mécanisme causant la dépression. L'hypothèse évolutionniste choisie par Wakefield (aussi défendue par l'Association américaine de psychiatrie⁵⁶) est celle du *breakdown*, selon laquelle les troubles sont dus à des dysfonctions des mécanismes mentaux. Par exemple, la dépression serait causée par une dysfonction du mécanisme mental, c'est-à-dire une réponse du mécanisme dont l'effet n'est pas explicable par une augmentation des chances de survie et de reproduction.

L'hypothèse du *mismatch* développée par Nesse et Williams⁵⁷ soutiendrait plutôt que les mécanismes mentaux seraient adaptés à l'environnement préhistorique et que les troubles seraient causés par le changement d'environnement depuis la préhistoire. Le contexte aurait changé rapidement et les mécanismes mentaux ne se seraient pas adaptés assez rapidement. Il en découle que les

états jugés pathologiques seraient le résultat de fonctions naturelles plutôt que de dysfonctions. Par exemple, la dépression serait issue d'un mécanisme naturel de conservation de l'énergie qui permettrait aux individus préhistoriques ayant accès à moins de ressources de survivre plus longtemps. Un tel mécanisme aurait contribué à la survie et à la reproduction des individus. Cette contribution en expliquerait la présence actuelle, même si ce mécanisme était jugé négativement aujourd'hui. Selon une telle explication évolutionniste, la fonction naturelle d'un mécanisme combinée au changement rapide d'environnement expliquerait la présence d'états pathologiques tels que l'état dépressif.

Selon l'hypothèse de la *persistance*, les troubles mentaux seraient occasionnés par des mécanismes mentaux qui suivent leur fonction naturelle même à notre époque⁵⁸. Ce serait le fonctionnement naturel d'un mécanisme qui expliquerait le mieux sa présence et qui serait la cause du préjudice⁵⁹. Comme dans l'hypothèse du *mismatch*, une fonction naturelle serait à l'origine du trouble mental, mais, ici, cette fonction contribuerait toujours à la survie et à la reproduction aujourd'hui. Par exemple, la dépression aurait comme fonction naturelle de permettre la « gestion des situations hostiles où la fuite est impossible⁶⁰ ». L'hypothèse de la *persistance* se démarque en maintenant que le mécanisme est toujours adaptatif de nos jours, ce qui signifie que le mécanisme suit sa fonction naturelle lorsqu'il est activé. En ce sens, la meilleure explication de la présence d'un mécanisme causant l'état dépressif serait sa fonction naturelle de gestion des situations hostiles.

Une réponse démesurée du mécanisme, la conservation de l'énergie et la gestion des situations hostiles seraient trois explications également plausibles de la présence d'un tel mécanisme. Puisque ces trois hypothèses seraient plausibles et qu'il n'y a que peu de connaissances sur les mécanismes mentaux, il ne serait pas possible de déterminer la meilleure explication⁶¹. L'indétermination du critère explicatif découle de cette coexistence : puisqu'il est impossible de déterminer précisément les fonctions des mécanismes mentaux, il n'est pas possible de savoir quels états

remplissent le critère explicatif de la HDA. Le critère explicatif est alors toujours indéterminé et inutilisable.

Wakefield juge toutefois que l'absence de connaissances précises ne rend pas les explications fonctionnelles inutiles⁶². Il juge que de telles explications sont utiles malgré l'absence d'une connaissance précise des fonctions des mécanismes. Il croit qu'il sera possible d'identifier, éventuellement, les mécanismes mentaux qui sont encore inconnus et demeure toujours optimiste quant cette possibilité. Nous croyons à l'inverse que l'imprécision des connaissances peut être la source de torts pour les individus. En ce sens, l'absence de connaissances précises limiterait l'action et les intuitions ne seraient pas suffisantes. Par exemple, sans connaissance des mécanismes fonctionnels de la fatigue, il serait possible de «soigner» la fatigue par une médication : l'individu ne subirait plus les conséquences de la fatigue, mais développerait probablement des conséquences bien plus graves⁶³. De la même manière, il est possible que des états normaux soient jugés dysfonctionnels en vertu d'une connaissance imprécise des mécanismes mentaux. De telles erreurs mèneraient à des conséquences dommageables pour les individus. Nous croyons donc, par opposition à Wakefield, que les explications évolutionnistes ne sont pas utiles lorsque jumelées à des connaissances imprécises des mécanismes.

Nous sommes ainsi amenés à conclure que le critère explicatif demeure nécessairement indéterminé. Nous avons présenté trois explications évolutionnistes concurrentes et plausibles. L'impossibilité de déterminer précisément laquelle de ces explications est applicable rendrait la clause factuelle indéterminée. En ce sens, une telle clause serait inutilisable, ne permettant pas à la HDA d'être véritablement utile.

Conclusion

Nous avons d'abord montré que l'intention initiale de laquelle est née la HDA était d'éviter les difficultés affectant les analyses normativiste et naturaliste. L'approche hybride devait fournir une troisième alternative. Nous avons ensuite montré que cette

alternative était problématique puisque le critère explicatif de la HDA s'appuie sur une thèse évolutionniste qui fait elle-même appel à des valeurs. En vertu de cette thèse, le critère évaluatif présente deux difficultés. Nous avons montré que la première concerne l'impossibilité de définir la dysfonction en termes factuels. Alors que la première formulation était explicitement appuyée sur des termes évaluatifs, la notion centrale d'« explication » de la reformulation serait elle aussi évaluative. Par conséquent, la HDA s'appuiera sur deux critères évaluatifs et rencontre les difficultés typiques des approches normativistes. De l'autre côté, trois explications évolutionnistes concurrentes de la présence des mécanismes mentaux rendent le critère explicatif constamment indéterminé. Il n'est alors pas possible de déterminer à quel moment le critère explicatif est rempli, ce qui rend ainsi la HDA inutilisable en pratique.

Puisque la thèse évolutionniste est à la source de ces deux difficultés, notre proposition consiste à remplacer le concept évolutionniste de fonction (au sens d'effet sélectionné) par un concept systémique. Cette proposition s'inspire du travail de Forest⁶⁴ qui constate que la conception systémique de la fonction a été écartée trop rapidement et mérite un examen plus approfondi. Selon lui, les médecins (et psychiatres) s'intéressent aux capacités des individus plutôt qu'aux mécanismes sélectionnés par l'évolution. Une conception systémique telle que développée par Robert Cummins⁶⁵ pourrait permettre de mieux rendre compte de cette intuition médicale⁶⁶.

1. Cf. Jerome C. Wakefield, « Disorder as Harmful Dysfunction : A Conceptual Critique of DSM-III-R 's Definition of Mental Disorder. », dans *Psychological Review*, vol. 99, n° 2, 1992, p. 232-247.
« Le concept de trouble mental : à la frontière entre faits biologiques et valeurs sociales », dans É. Giroux et M. Lemoine (dir.), *Philosophie de la médecine : santé, maladie, pathologie*, trad. Steeves Demazeux, Vrin, 2012, p. 127-176.

2. Christopher Boorse, «Le concept théorique de santé», dans É. Giroux et M. Lemoine (dir.), *Philosophie de la médecine : santé, maladie, pathologie*, trad. Élodie Giroux, Vrin, 2012, p. 61-119.
3. Robert Evan Kendell, «The concept of disease and its implications for psychiatry.», dans *The British journal of psychiatry: the journal of mental science*, n° 127, 1975, p. 305-315.
4. Kenneth C. Waters, «What Concept Analysis in Philosophy of Science Should Be (and Why Competing Philosophical Analyses of Gene Concepts Cannot Be Tested by Polling Scientists).», dans *History and Philosophy of the Life Sciences*, vol. 26, n° 1, 2004, p. 29.
5. Son efficacité n'est toutefois pas entièrement acceptée en philosophie de la psychiatrie. Cf. Rachel Valerie Cooper, *Psychiatry and philosophy of science*, Montreal, McGill-Queen's University Press, 2007. Maël Lemoine, «Defining disease beyond conceptual analysis: an analysis of conceptual analysis in philosophy of medicine.», dans *Theoretical Medicine and Bioethics*, vol. 34, n° 4, 2013, p. 309-325.
6. Forest et Faucher affirment que «Jerome Wakefield's work is at the center of the contemporary debate as to the nature of mental illness (and the related question of psychiatry's scope and limits), a decades-old debate in both scientific and philosophical literature. His key proposal, the "harmful dysfunction analysis" of mental disorders (HDA thereafter), has been greatly discussed by scientists and philosophers alike.» Cf. Denis Forest et Luc Faucher. «Introduction», dans L. Forest & L. Faucher (dir.), *Defining mental disorder*, Cambridge, MIT Press, 2021, p. ix.
7. Denis Forest, «De quel concept de fonction la philosophie de la médecine peut-elle avoir besoin?», dans *Revue philosophique de la France et de l'étranger*, vol. 134, n° 1, 2009, p. 59.
8. Jerome C. Wakefield, «Fait et valeur dans le concept de trouble mental: le trouble en tant que dysfonction préjudiciable.», dans *Philosophiques*, trad. Sébastien Cloutier, vol 33, n° 1, 2006, p. 43.
9. Marc Ereshefsky, «Defining "health" and "disease".», dans *Studies in history and philosophy of biological and biomedical sciences*, vol. 40, n° 3, 2009, p. 222.
10. Christopher Boorse, «On the Distinction between Disease and Illness.», dans *Philosophy & Public Affairs*, vol. 5, n° 1, 1975, p. 49-68.

11. Christopher Boorse, «Le concept théorique de santé», *loc. cit.*, p. 62.
12. Peter Sedgwick, *Psychopolitics*, New York, Harper and Row, 1982, p. 32, nous traduisons.
13. Scott O. Lilienfeld et Lori Marino, «Mental Disorder as a Roschian Concept: A Critique of Wakefield's "Harmful Dysfunction" Analysis.», dans *Journal of Abnormal Psychology*, vol. 104, n° 3, 1995, p. 411.
14. Steeves Demazeux, «Présentation du texte de Jerome Wakefield», dans É. Giroux et M. Lemoine (dir.), *Philosophie de la médecine : santé, maladie, pathologie*, Vrin, 2012, p. 122.
15. Jerome C. Wakefield, «Fait et valeur dans le concept de trouble mental: le trouble en tant que dysfonction préjudiciable.», *loc. cit.* p. 42.
16. *Ibid.*, p. 41.
17. Scott O. Lilienfeld et Lori Marino, «Mental Disorder as a Roschian Concept: A Critique of Wakefield's "Harmful Dysfunction" Analysis.», dans *Journal of Abnormal Psychology*, vol. 104, n° 3, 1995, p. 411-420.
18. *Ibid.*, p. 416.
19. Rachel Valerie Cooper, *op. cit.*, p. 39.
20. Lawrie Reznek. *The nature of disease*, New York, Routledge & Kegan Paul, 1987, p. 163-164.
21. Jerome C. Wakefield, «Le concept de trouble mental : à la frontière entre faits biologiques et valeurs sociales», *loc. cit.*, p. 169.
22. Cf. Jerome C. Wakefield, «Fait et valeur dans le concept de trouble mental: le trouble en tant que dysfonction préjudiciable.», *loc. cit.*
23. *Ibid.*, p. 374.
24. Cf. Jerome C. Wakefield et Jordan A. Conrad, «Does the harm component of the harmful dysfunction analysis need rethinking?: Reply to Powell and Scarffe.», dans *Journal of Medical Ethics*, vol. 45, n° 9, 2019, p. 594-596.
25. Karen Neander, «The Teleological Notion of "Function.», dans *Australasian Journal of Philosophy*, vol. 69, n° 4, 1991, p. 454-468.
26. Jerome C. Wakefield et Jordan A. Conrad, *loc. cit.*, p. 594-596.
27. Jerome C. Wakefield. «The concept of mental disorder: diagnostic implications of the harmful dysfunction analysis.», dans *World psychiatry: official journal of the World Psychiatric Association (WPA)*, vol. 6, n° 3, 2007, p.152.

28. *Ibid.*, p. 152
29. Jerome C. Wakefield, «Dysfunction as a Value-Free Concept : A Reply to Sadler and Agich.», dans *Philosophy, Psychiatry, & Psychology* 2, n° 3, 1995, p. 235.
30. Jerome C. Wakefield, «Fait et valeur dans le concept de trouble mental.», *loc. cit.*, p. 47.
31. *Ibid.*, p. 49.
32. Jerome C. Wakefield, «Le concept de trouble mental : à la frontière entre faits biologiques et valeurs sociales», *loc. cit.*, p. 165.
33. David J. Buller. *Adapting minds : evolutionary psychology and the persistent quest for human nature*. Cambridge, The MIT Press, 2005, p. 69, nous traduisons.
34. Luc Faucher. «Facts, Facts, Facts : HD Analysis Goes Factuel», dans L. Forest & L. Faucher (dir.), *Defining mental disorder*, Cambridge, MIT Press, 2021, p. 63.
35. Cela constitue toutefois un présupposé questionné. Cf. Anne Gammelgaard, «Evolutionary biology and the concept of disease.», dans *Medicine, Health Care and Philosophy*, vol. 3, n° 2, 2000, p. 110.
36. Jerome C. Wakefield, «Fait et valeur dans le concept de trouble mental.», *loc. cit.*, p. 43.
37. Rachel Valerie Cooper, *Op. cit.*, p. 43. Steeves Demazeux, «The Function Debate and the Concept of Mental Disorder.», dans *Classification, Disease and Evidence : New Essays in the Philosophy of Medicine*, Dordrecht, Springer Netherlands, 2015, p. 79. Pierre-Henri Castel, «Toward a More Constructive (and Historical) View of the Harmful Dysfunction Theory?», dans *Philosophy, Psychiatry, & Psychology*, vol. 21, n° 4, 2014, p. 358.
38. Jerome C. Wakefield, «Le concept de trouble mental : à la frontière entre faits biologiques et valeurs sociales», *loc. cit.*, p. 39.
39. Cf. Karen Neander, «The Teleological Notion of “Function”.», *loc. cit.* «Les explications fonctionnelles.», dans *Revue philosophique de la France et de l'étranger*, vol. 134, n° 1, 2009, p. 5-34. «Functional analysis and the species design.», dans *Synthese*, vol. 194, n° 4, 2017, p. 1147-1168.
40. Le projet descriptif est remis en question autant par les défenseurs que par les critiques de la conception évolutionniste de «dysfonction» et de la HDA. Cf. Paul E. Griffiths et John

- Matthewson. «Evolution, Dysfunction, and Disease : A Reappraisal.», dans *British Journal for the Philosophy of Science*, vol. 69, n° 2, 2018, p. 301-27. Maël Lemoine. «Is the Dysfunction Component of the “Harmful Dysfunction Analysis” Stipulative?», dans L. Forest & L. Faucher (dir.), *Defining mental disorder*, Cambridge, MIT Press, 2021, p. 199-212.
41. John Z. Sadler et George J. Agich, «Diseases, Functions, Values, and Psychiatric Classification.» dans *Philosophy, Psychiatry, & Psychology*, vol. 2, n° 3, 1995, p. 219.
 42. *Ibid.*, p. 223.
 43. Jerome C. Wakefield, «Disorder as Harmful Dysfunction», *loc. cit.*, p. 237.
 44. Anne Gammelgaard, *loc. cit.*, p. 109.
 45. Jerome C. Wakefield, «Le concept de trouble mental : à la frontière entre faits biologiques et valeurs sociales», *loc. cit.*, p. 162. Nous soulignons.
 46. John Z. Sadler, et George J. Agich, *loc. cit.*, p. 219-231.
 47. Jerome C. Wakefield, «Dysfunction as a Value-Free Concept : A Reply to Sadler and Agich.», *loc. cit.*, p. 234.
 48. Jerome C. Wakefield, «Le concept de trouble mental : à la frontière entre faits biologiques et valeurs sociales», *loc. cit.*, p. 50.
 49. Scott DeVito, «On the Value-Neutrality of the Concepts of Health and Disease : Unto the Breach Again.», dans *The Journal of Medicine and Philosophy*, vol. 25, n° 5, 2000, p. 554.
 50. Griffiths et Matthewson, par exemple, présentent de bonnes raisons de défendre une conception évolutionniste de la dysfonction. Même s'ils évaluent un bon nombre de critiques de la conception évolutionniste de la fonction, la remise en question de son caractère factuel n'en fait pas partie. Appuyer le concept de fonction sur une «science mature» serait suffisant pour que le concept de dysfonction soit factuel. Nous ne partageons pas un tel présupposé. Cf. Paul E. Griffiths et John Matthewson. *loc. cit.*
 51. Anne Gammelgaard, *loc. cit.*, p. 114.
 52. L'on pourrait aussi défendre qu'un rapport à des buts est toujours évaluatif, montrant ainsi que le critère explicatif n'est pas véritablement factuel. Cf. Christopher Megone, «Mental Illness, Human Function, and Values.», dans *Philosophy, Psychiatry, & Psychology*, vol. 7, n° 1, 2000, p. 55.

53. Marc Ereshefsky, *loc. cit.*, p. 223.
54. *Ibid.*, p. 225.
55. Somogy Varga, «Defining mental disorder. Exploring the ‘natural function’ approach.», dans *Philosophy, Ethics, and Humanities in Medicine*, vol. 6, 2011, p. 1-10. Dominic Murphy, «Can Evolution Explain Insanity?», dans *Biology and Philosophy*, vol. 20, n° 4, 2005, p. 745-766.
56. Rachel Valerie Cooper, *op. cit.*, p. 33.
57. Somogy Varga, *loc. cit.*, p. 6.
58. Cf. Anne Gammelgaard, *loc. cit.*, p. 109-116. Randolph M. Nesse, «Natural selection and the regulation of defenses : A signal detection analysis of the smoke detector principle.», dans *Evolution and Human Behavior*, vol. 26, n° 1, 2005, p. 88-105. «Darwinian medicine and mental disorders.», dans *International Congress Series*, vol. 1296, n° C, 2006, p. 83-94.
59. Dominic Murphy et Robert L. Woolfolk, «The Harmful Dysfunction Analysis of Mental Disorder.» dans *Philosophy, Psychiatry, & Psychology*, vol. 7, n° 4, 2000, p. 241-52.
60. Somogy Varga, *loc. cit.*, p. 6.
61. Jonathan Y. Tsou, «Psychiatry in the Scientific Image», dans *Philosophical Psychology*, vol. 21, n° 1, 2008, p. 136.
62. Jerome C. Wakefield, «The concept of mental disorder : diagnostic implications of the harmful dysfunction analysis.», *loc. cit.*, p. 151. «Fait et valeur dans le concept de trouble mental.», *loc. cit.*, p. 47.
63. Randolph M. Nesse, «Darwinian medicine and mental disorders.», *loc. cit.*, p. 87.
64. Denis Forest, «De quel concept de fonction la philosophie de la médecine peut-elle avoir besoin?», dans *Revue philosophique de la France et de l'étranger*, vol. 134, n° 1, 2009, p. 64.
65. Robert Cummins, «Functional Analysis.», dans *The Journal of Philosophy*, vol. 72, n° 20, 1975, p. 758.
66. Je souhaite remercier Marianne Vigneault, le comité de révision de la revue *Phares* et Pierre-Olivier Méthot pour leurs commentaires sur les différentes versions de ce texte.