

Quel cadre heuristique la philosophie analytique des émotions peut-elle apporter au rôle des états affectifs dans les modèles de la dépression ?

Le conflit des passions et de la raison, qui mener jusqu'à la folie, est le motif ancien qui est encore à l'horizon de la psychiatrie lorsque celle-ci apparaît en Europe au XIX^e siècle : ce thème philosophique sous-tend l'approche médicale de la folie inaugurée par les aliénistes. La mélancolie est d'abord vue comme un délire causé par des passions tristes, puis Jules Séglas décrit une mélancolie sans délire qui préfigure la dépression d'aujourd'hui (Haustgen, 2014). Celle-ci est vue comme un trouble affectif, caractérisé par des sentiments de tristesse et d'appréhension (Radden, 2009), qui va entrer dans la nosologie comme un trouble de l'humeur. Aujourd'hui la dépression est définie à partir de l'épisode dépressif caractérisé, défini dans les classifications (le DSM, manuel diagnostique et statistique des troubles mentaux, et la CIM, classification internationale des maladies) à partir de deux symptômes centraux qui sont l'humeur dépressive et la perte d'intérêt et de plaisir, dont un des deux au moins doit être présent, ainsi que sept symptômes secondaires dont quatre au moins doivent être présents en association avec l'un des deux symptômes centraux pendant une durée d'au moins deux semaines, et représenter un changement significatif du fonctionnement habituel du patient en association avec une souffrance cliniquement significative. Les symptômes secondaires sont les troubles du sommeil, de l'appétit, la perte d'énergie, les difficultés de concentration, les sentiments de culpabilité ou de dévalorisation excessive, le fait de penser à la mort ou au suicide (*American psychiatric association*, 2013). La clinique fait des distinctions concernant d'une part l'épisode dépressif (sévérité, chronicité), d'autre part les patients (réponse au traitement, en rémission complète ou partielle, avec des symptômes résiduels, guérison, rechute), pour s'orienter entre les traitements qui sont les psychothérapies (psychodynamiques, psychanalytiques, thérapies cognitivo-comportementales), les stratégies médicamenteuses (plusieurs générations d'anti-dépresseurs, anxiolytiques), ou alors l'hospitalisation et les traitements de dernier recours (électroconvulsivo-thérapie, perfusions de kétamine) dans le cas des dépressions résistantes où le patient ne répond pas aux autres traitements. Ces distinctions orientent les recherches des sciences cliniques pour caractériser la prévalence de la dépression dans la population clinique et non clinique, rechercher des facteurs de la dépression, ou encore chercher à distinguer des sous-types de dépressions. La recherche clinique s'interroge particulièrement sur les facteurs de récurrence de l'épisode dépressif caractérisé, car c'est la répétition de celui-ci chez un même patient qui permet de parler chez ce dernier d'un « trouble dépressif unipolaire », parfois appelé « maladie dépressive », qui est la dépression au sens d'un trouble chronique et invalidante, même si l'absence d'une étiologie connue rend le terme de « maladie » incertain (Corruble, 2010).

Le point de départ des sciences cliniques est la symptomatologie qui est à la fois affective, somatique et cognitive et comportementale : si l'humeur dépressive et les perturbations affectives allant d'une atténuation de l'ensemble des états affectifs à des émotions négatives très marquées, tristesse, peur, mais aussi colère, honte ou culpabilité, sont centrales, la symptomatologie dépressive est également cognitive, avec par exemple des ruminations et des troubles de la concentration, comportementale avec une baisse générale de l'activité et de la motivation, et somatique avec en particulier des troubles de l'appétit et du sommeil.

En l'absence d'un modèle explicatif unifié des troubles mentaux, et d'une étiologie connue à la dépression, une variété de modèles ont été proposés, dont certains donnent une place centrale aux états affectifs. Ce qui fait défaut en l'absence d'un modèle explicatif étiologique, c'est un critère du pathologique dans les troubles mentaux : une approche évolutionniste offre alors un cadre heuristique propice à leur interprétation, évaluant les traits psychologiques et comportementaux en jeu dans les tableaux cliniques à l'aune de leurs fonctions adaptatives, donc des effets en vue desquels ils auraient été retenus par la sélection naturelle, suivant les principes de la « médecine darwinienne » (Nesse et al., 1995; Nesse & Williams, 1998). Ces approches substituant la question « pourquoi la dépression ? » à la question « comment la dépression ? » ont donné lieu à des modèles

interprétant la dépression par des états affectifs ne remplissant plus leur fonction évolutive (Nesse, 2009). Mais ces modèles, dont aucun ne parvient à s'imposer, rencontrent des limites et sont soumis à des critiques (Faucher, 2016).

Puisque la question est alors de savoir si les états affectifs, en particulier les émotions et les humeurs, ont une fonction évolutive et si oui laquelle, la philosophie des émotions peut-elle venir en aide à ces modèles ? Une théorie évolutionniste des émotions répondrait-elle à cette question et soutiendrait-elle dès lors ce type de modèle évolutionniste de la dépression ? Dans la mesure où la possibilité d'une explication évolutionniste des émotions est autant critiquée au sein de la philosophie des émotions que les approches évolutionnistes le sont au sein des modèles de la dépression, y a-t-il sinon d'autres théories des émotions qui pourraient fournir d'autres perspectives sur les perturbations affectives propres à la dépression ? Dans l'attente d'une découverte des causes de la dépression et d'un modèle explicatif de celle-ci par les sciences de base, la philosophie des émotions peut-elle apporter un cadre heuristique non évolutionniste à des mécanismes affectifs pathologiques caractéristiques de ce trouble ?

Dans une première partie je propose d'étudier les apports et les limites d'un modèle évolutionniste de la dépression qui l'interprète comme une émotion pathologique au sens d'une émotion ne remplissant plus sa fonction adaptative, en mettant en évidence des limites de fait d'un tel modèle mais aussi une limite de principe, qui se retrouve dans une théorie évolutionniste des émotions et l'empêche de venir en appui à un tel modèle. Dans une deuxième partie, je dégage de la littérature une caractérisation non évolutionniste des perturbations des états affectifs caractéristiques de la dépression, et j'identifie trois cadres heuristiques possibles pour ces perturbations à partir des différentes théories présentes dans la philosophie des émotions.

1. Une théorie évolutionniste des émotions peut-elle soutenir un modèle évolutionniste de la dépression comme émotion pathologique ?

1. 1. D'un modèle évolutionniste de la dépression à un modèle évolutionniste de la dépression comme émotion pathologique

Les modèles évolutionnistes de la dépression envisagent celle-ci comme une réponse comportementale : ils rapportent les différents aspects de la dépression à une fonction évolutive consistant dans le fait qu'elle est un trait comportemental qui répondait à certains enjeux de l'environnement dans lequel vivaient nos ancêtres, chez qui elle a représenté une adaptation. Autrement dit la dépression serait un trait comportemental qui aurait été naturellement sélectionné car il était avantageux pour les individus qui le présentaient. Cet environnement est dans ces différents modèles l'EEA (« environnement of evolutionary adaptedness ») qui a été théorisé par Tooby et Cosmides comme un environnement proche du pléistocène, période au cours de laquelle la plupart des adaptations se seraient accomplies (Barkow et al., 1992).

L'enjeu d'un modèle évolutionniste de la dépression est alors premièrement d'expliquer à quel type de problème de cet environnement la dépression était susceptible de répondre, deuxièmement de rendre compte de ses différents aspects relativement à cette fonction, et troisièmement de déterminer si dans notre environnement actuel elle remplit toujours sa fonction ou bien si, cet environnement étant trop différent de l'EEA, elle constitue désormais une réponse inadaptée (Mouchabac, 2006).

Par exemple le modèle de la compétition sociale (Price et al., 1994) explique la dépression comme une stratégie d'acceptation de la défaite au sein d'un groupe social, par laquelle l'individu qui n'est pas en mesure de remporter la compétition se met lui-même hors jeu en adoptant un état dans lequel il n'est plus motivé à contre-attaquer, signale à ses partenaires ou adjuvants de ne plus compter sur lui pour combattre, et montre à ses adversaires qu'il se retire de la lutte. Dans ce modèle la dépression est une réponse comportementale fonctionnelle, c'est-à-dire remplissant toujours sa fonction évolutive, et dont la prévalence de plus en plus forte serait à mettre au compte d'un environnement qui multiplie désormais les situations de compétition sociale.

À l'inverse dans le modèle de l'humeur basse (Nesse, 2009) la dépression est dans de nombreux cas une réponse dysfonctionnelle, consistant dans une humeur basse qui ne remplit plus sa fonction évolutive. L'humeur basse est un ensemble de modifications concernant la physiologie, la motivation et le comportement, visant à réguler le comportement par rapport aux enjeux d'une situation pour l'organisme, plus précisément déclenché par des situations peu avantageuses pour l'organisme dans lesquelles il est plus profitable pour ce dernier de réduire son activité pour ne pas prendre de risque (Nesse, 2009). L'humeur basse possède une fonction, elle est adaptative, et avec elle la dépression dans certains cas. Mais elle peut se déclencher chez certains individus en l'absence des contraintes environnementales qui la justifient du point de vue de l'évolution, en raison de défaillances des mécanismes qui la sous-tendent au niveau biologique : dans ce cas la dépression est une humeur basse pathologique en raison de désajustements entre l'EEA et notre environnement actuel, ou bien d'un dysfonctionnement intrinsèque.

Enfin certains modèles prennent en compte la variabilité interindividuelle des traits retenus au sein d'une population et expliquent la dépression non comme une adaptation mais comme une réponse comportementale présente chez certains individus parce que ceux-ci sont porteurs de traits à un degré extrême qui les rend vulnérables à cette réponse comportementale dans certaines situations : c'est le cas du modèle de Nettle dans lequel le trait comportemental en question est le neuroticisme, le fait d'être plus ou moins réactif émotionnellement (Nettle, 2004).

Mais le fait d'aborder la dépression comme une réponse comportementale peut être vu comme une limite de ces modèles, même lorsqu'ils rendent possible de la concevoir comme une

réponse dysfonctionnelle, parce que définir la dépression sur le plan comportemental les empêche de rendre compte de ce qu'il peut y avoir de pathologique dans le vécu de la dépression en termes de souffrance et donc de préjudice subi. Si aucun de ces modèles évolutionnistes ne peut pour l'instant être mis en cohérence avec la clinique, c'est parce qu'ils ne font pas de véritable place à la détresse subjective des patients dépressifs, alors que c'est cette détresse qui est invalidante et pousse les patients à consulter un médecin : c'est ce qui laisse les cliniciens peu convaincus par ces modèles (Mouchabac, 2006).

D'une part la réponse comportementale dépressive, si elle remplit une fonction, doit être avantageuse pour l'individu, représenter pour lui la solution d'un problème, or tel n'est pas le cas. Néanmoins certains de ces modèles peuvent rendre compte de cela : ou bien par la thèse d'une modification de notre environnement par rapport à l'EEA qui crée un désajustement entre nos traits comportementaux et les contraintes de l'environnement actuel, ou bien par la thèse d'une variabilité interindividuelle de certains traits comportementaux dans la population. Mais d'autre part ces modèles évolutionnistes ne rendent pas bien compte du vécu affectif de la dépression : puisque la dépression est traitée comme une réponse comportementale, ce vécu affectif est simplement considéré comme une cause ou une conséquence de cette réponse comportementale, ou bien son corrélat, son épiphénomène au statut ontologique incertain. Dans le modèle de la compétition sociale, en termes de vécu affectif, une humeur dépressive est mentionnée comme ce qui entraîne la réponse comportementale dépressive : elle fait la médiation entre l'information reçue par l'individu sur sa défaite dans la compétition sociale et sa variation comportementale (Price et al., 1994). Dans le modèle de l'humeur basse, cette humeur semble consister dans l'expérience de la réponse comportementale elle-même, implicitement réductible à celle-ci puisque Nesse définit les émotions et les humeurs comme un ensemble de modifications de la physiologie, de la motivation et du comportement (Nesse, 2009). Mais il semble difficile de réduire le vécu affectif de la dépression à l'expérience de cette réponse comportementale : Nesse reconnaît qu'il y a aussi dans la dépression des affects négatifs d'intensité sévère distincts de l'humeur basse (Mouchabac, 2006).

Finalement le vécu affectif dépressif ne saurait être jugé proportionnel à la fonction supposée de la réponse comportementale qu'il accompagne, et semble phénoménologiquement différent de tout vécu affectif normal (Faucher, 2016). C'est en définitive ce qui fait sérieusement douter de la possibilité que cette réponse comportementale elle-même soit passible d'une explication évolutionniste la rapportant directement ou indirectement à des traits adaptatifs (remplissant une fonction évolutive). Ce vécu affectif pourrait-il n'être qu'un trait secondaire de la réponse comportementale, sélectionné de façon liée à cette réponse, et non parce qu'il procure lui-même un avantage ? Mais on pourrait alors se demander si le préjudice subi dans ce vécu affectif, et le retentissement dont ce préjudice s'accompagne pour l'individu, soit le degré d'incapacité qu'entraîne la souffrance affective de la dépression, qui conduit notamment au suicide, est compatible avec sa conservation au cours de l'histoire évolutive. Peut-on maintenir que la réponse comportementale résulte de traits présentant des avantages suffisants pour compenser sa létalité, pour que son maintien s'explique dans la dynamique des populations modélisable par la phylogénie, ou que celui-ci puisse être un hasard stochastique ?

Si on laisse cette question en attente de résultats scientifiques ultérieurs, l'alternative pour contourner ce problème dès maintenant pourrait être de faire un modèle évolutionniste de la dépression qui partirait du vécu affectif lui-même plutôt que du comportement ; autrement dit caractériser la dépression comme une expérience singulière dans la vie affective, à partir de laquelle il faudrait rendre compte de ses traits comportementaux, plutôt que comme un type de comportement spécifique dont il faudrait ensuite justifier le vécu affectif. Puisqu'il semble difficile ou impossible de conférer un avantage adaptatif au vécu affectif de la dépression, un tel modèle devrait l'expliquer comme dysfonctionnel. La dépression serait un vécu affectif dysfonctionnel au sens d'un ensemble d'états affectifs ne remplissant plus leur fonction adaptative. C'est l'idée d'un modèle évolutionniste de la dépression comme émotion pathologique.

1. 2. Un modèle évolutionniste de la dépression comme émotion pathologique et ses limites

Un modèle évolutionniste de la dépression comme émotion pathologique repose sur la thèse évolutionniste selon laquelle il est possible de rendre compte des émotions d'un point de vue évolutif. Cela requiert de déterminer premièrement une fonction adaptative des émotions, un « effet sélectionné », quelque chose que les émotions permettent et qui est avantageux pour leurs porteurs ce qui en fait des traits retenus dans la population par la sélection naturelle (Millikan, 1989; Neander, 1991), et deuxièmement cela requiert de rendre compte de leurs caractéristiques relativement à cette fonction. Dans un tel modèle la dépression est une émotion dysfonctionnelle, au sens où elle ne remplit plus sa fonction évolutive, et même pathologique dans la mesure où ce dysfonctionnement s'accompagne d'un préjudice (Wakefield, 1992).

Le principal modèle de ce type-là est le modèle de la « tristesse pathologique » (Horwitz & Wakefield, 2007). La dépression est alors une tristesse qui ne remplit plus sa fonction évolutive, le rôle pour lequel la sélection naturelle l'a retenue, qui est de permettre la réparation d'une perte, parce qu'elle intervient dans une situation où aucune perte n'a été subie, ou bien de façon disproportionnée par rapport à la perte subie. Le cadre évolutionniste est déterminant : la tristesse est pathologique parce qu'elle porte préjudice à l'individu qui en souffre, mais uniquement dans la mesure où elle résulte d'un dysfonctionnement relativement à une fonction évolutive, comme l'a réaffirmé Wakefield en réponse aux critiques de son modèle du trouble mental (Faucher & Forest, 2021).

Horwitz et Wakefield développent donc une « anatomie de la tristesse normale » qui en est une approche évolutionniste visant principalement à montrer qu'elle constitue une adaptation, un trait psychologique retenu par la sélection naturelle car il est avantageux pour ses porteurs : ils déterminent donc une fonction évolutive de la tristesse et rendent compte de ses caractéristiques relativement à celle-ci (Horwitz & Wakefield, 2007). Ils s'appuient sur différentes littératures empiriques pour montrer que la tristesse normale est un trait qui a pour fonction de permettre la réparation d'une perte. Étant un trait normal du « design » de l'être humain, elle est donc liée à un contexte qui la cause, proportionnelle à cette cause en intensité et s'arrête lorsqu'elle a permis la réparation de la perte subie. Son contexte est une situation typique de perte : deuil, problèmes dans les relations sociales, perte d'un emploi ou d'un statut social, stress chronique, désastre subi... La justification de la possibilité d'une approche évolutionniste de la tristesse, est donnée par la littérature empirique qui montre 1) que la tristesse est un trait comportemental commun à l'être humain et à d'autres espèces animales, 2) qu'elle est présente chez l'enfant pré-socialisé, 3) qu'elle est uniforme à travers les différentes cultures. Une littérature empirique montre qu'une part de la tristesse normale est biologique, une autre part est culture-dépendante : la culture influence la définition de ce qui correspond à une perte et la façon dont il convient d'y répondre, d'où une relativité culturelle de la démarcation entre une tristesse fonctionnelle, qui remplit sa fonction évolutive, et une tristesse dysfonctionnelle, qui ne remplit pas cette fonction, et peut dès lors être pathologique. Ainsi le taux de dépression peut varier selon les aires culturelles sans que cela empêche de la définir de façon pan-culturelle comme une tristesse dysfonctionnelle, ne remplissant plus sa fonction évolutive de réparation d'une perte. Une littérature empirique rend enfin compte de différents mécanismes par lesquels la tristesse peut permettre à un individu qui a subi une perte d'y répondre : attirer le soutien social qui va pouvoir contribuer à la réparation de la perte, protéger d'une nouvelle agression après une perte déjà subie en démotivant des comportements agressifs de la part des autres, se désengager d'activités non productives vis-à-vis de la réparation de la perte pour se « concentrer » sur celle-ci. Tous ces éléments soutiennent la thèse d'une fonction évolutive de la tristesse comme réponse à la perte.

D'après les auteurs, seule cette fonction évolutive de réponse à la perte permet de rendre compte de la tristesse, avec tous ses traits (Horwitz & Wakefield, 2007). Ils pointent « l'échec des sciences sociales pour distinguer la tristesse du trouble dépressif » : ils affirment que les approches non évolutionnistes de la tristesse, qui l'expliquent donc comme un trait culturel historiquement déterminé, en anthropologie et en sociologie, ne peuvent pas expliquer qu'en dépit de formes culturellement variables, celle-ci se prête à un type de dysfonctionnement pan-culturel. Le défaut des modèles issus de ces sciences sociales serait de s'en tenir à une compréhension « locale » de ce qu'est la tristesse : ils en décrivent les formes concrètes culturellement variables, du point de vue desquelles tristesse normale et pathologique peuvent être indiscernables. Rendre compte du pathologique, supposerait de transcender la part culturellement variable des émotions en identifiant ce qui en fait une caractéristique pan-culturelle du « design » humain normal, donc une approche évolutionniste (Horwitz & Wakefield, 2007).

Apports de ce modèle

Ce modèle explicatif de la dépression accorde une place centrale au vécu affectif, en cohérence avec la définition de ce trouble dans la nosologie comme en psychopathologie et dans la clinique, en rendant compte de ce vécu lui-même comme pathologique. Ce modèle a des apports pour la clinique, pour la nosologie, pour les modèles explicatifs et pour la définition de la dépression en tant que trouble mental.

1) Le bénéfice pour la clinique est que le modèle fournit une définition du fonctionnement psychologique normal, concernant le trait central de la dépression qu'est le vécu affectif. Ce fonctionnement normal consiste dans le fait que les traits de ce vécu affectif, en l'occurrence les émotions, remplissent leurs fonctions évolutives. En supposant que la science, en l'occurrence la psychologie évolutionniste, puisse spécifier suffisamment ces fonctions et ce fonctionnement normal, cela donne une base objective à la notion de vécu affectif dysfonctionnel, donc une base objective au pathologique, contrairement à la seule prise en compte du préjudice subi dans ce vécu affectif douloureux, qui serait une base seulement subjective pour le pathologique. Or en santé mentale, une base objective du pathologique fait justement défaut, en l'absence d'un modèle étiologique unifié des troubles mentaux. Pour la dépression en particulier, cela pose problème pour la définition de l'épisode dépressif, en raison de la ressemblance entre ses manifestations extérieures et celles de certains épisodes de tristesse normale. Dans le DSM-5 (manuel diagnostique et statistique des troubles mentaux, cinquième édition) la définition de l'épisode dépressif par une liste de symptômes s'assortit d'une durée de deux semaines qui semble relativement arbitraire ; la précision selon laquelle il incombe au clinicien de déterminer si ces symptômes sont cliniquement significatifs laisse implicitement ouverte la question d'un éventuel cadre théorique auquel les cliniciens s'en remettraient pour évaluer cette significativité des manifestations cliniques observées. Le principal avantage du modèle de la tristesse pathologique est qu'il explicite ce cadre, et guide les cliniciens pour rapporter les symptômes au contexte dans lequel ils surviennent : au-delà de la souffrance du patient, il s'agit de déterminer si ses symptômes sont le signe d'une tristesse qui ne remplit plus sa fonction de permettre la réparation d'une perte (Horwitz & Wakefield, 2007).

2) Dans la nosologie, l'approche « descriptive » du DSM, reprise par la CIM (classification internationale des maladies) doit permettre à ces classifications de servir à la communication entre clinique et recherche de telle sorte que la recherche puisse trouver son point de départ dans la clinique et que la clinique puisse faire évoluer ses catégories diagnostiques en fonction des résultats de la recherche au fur et à mesure des éditions successives (Demazeux, 2013). Si aucun modèle étiologique ne s'est imposé à ce jour, l'orientation des recherches vers une étiologie biologique est claire notamment dans les projets de nouvelles classifications (Faucher & Goyer, 2015). Or le modèle de la tristesse pathologique s'accorde naturellement avec la recherche d'une étiologie biologique de la dépression, reposant sur la thèse selon laquelle il est possible de définir les

émotions par une théorie de psychologie évolutionniste naturaliste où elles sont en partie des réactions biologiques.

3) Concernant les modèles explicatifs de la dépression, même s'il était impossible d'aboutir à un modèle étiologique strictement biologique, un modèle évolutionniste permet l'articulation de modèles explicatifs partiels donnés par différentes sciences de base à différents niveaux explicatifs. On peut distinguer six familles de modèles : les modèles psychanalytiques qui expliquent la dépression par des mécanismes intrapsychiques, les modèles comportementaux et environnementaux qui l'expliquent par une extinction du renforcement positif du comportement, les modèles cognitifs qui l'expliquent par des mécanismes cognitifs défaillants, les modèles systémiques qui l'expliquent comme adaptation d'un système à une crise, système incluant l'individu et les groupes sociaux dont il fait partie, les modèles neurobiologiques qui mettent en évidence des défauts d'interconnexion neurale notamment entre des régions corticales et limbiques, et enfin les modèles dits éthologiques, dérivés de la psychologie évolutionniste, qui permettent l'intégration de données multifactorielles (Desseilles, 2010). Le modèle évolutionniste de la tristesse pathologique, peut intégrer les explications des autres modèles puisqu'il se fonde sur la thèse que l'émotion est passible d'une explication évolutionniste naturaliste, donc sur une théorie de celle-ci qui englobe tous ses aspects, cognitifs, comportementaux, neurobiologiques, sociaux et psychiques.

4) Au-delà de cette capacité d'intégration, un modèle évolutionniste de la dépression comme émotion pathologique a un apport définitionnel : alors que l'éclatement des modèles étiologiques de la dépression en une myriade de modèles explicatifs partiels, cherchant les mécanismes défaillants derrière tel ou tel de ses aspects cliniques, finit par faire perdre de vue ce qui la distingue d'une maladie « somatique », la définir comme une émotion pathologique permet de la définir comme un trouble spécifiquement mental. L'émotion étant un état mental, ou psychologique, avec une structure intentionnelle, un tel modèle confère en effet du même coup une structure intentionnelle à la dépression, et donc ramène sa définition sur le plan du mental. Si la dépression est une tristesse pathologique, elle a une intentionnalité qui consiste dans la représentation d'une perte. Cela lui donne un sens, présumé par la clinique puisque l'interprétation du tableau clinique et les approches thérapeutiques se déroulent et cherchent à agir sur le plan psychique. Pierre-Henri Castel affirme ainsi que c'est l'approche évolutionniste qui a contribué à re-mentaliser la dépression, même si elle s'inscrit dans une vision naturaliste de la texture intentionnelle morale de celle-ci qu'il n'estime pas tenable (Castel, 2009). Tout en étant une approche naturaliste de la dépression (supposant implicitement que l'intentionnalité de l'émotion pourrait être naturalisée), le modèle de la tristesse pathologique tend donc à redéfinir la dépression comme un trouble mental, ce qui a trois avantages sur le plan définitionnel.

- Premièrement, le modèle de la tristesse pathologique donne un sens unifié à la dépression, ce qui le rend plus satisfaisant que les modèles explicatifs partiels. Alors que la souffrance induite par la maladie mentale est d'autant plus problématique qu'elle est inintelligible, un tel modèle lui confère une signification : avoir en tête que la dépression est une forme de tristesse, donc représente une perte, ouvre une piste à la compréhension de telle dépression particulière, de son sens, véritable moyen d'action dans l'ordre du psychique.

- Deuxièmement ce modèle cherche à expliquer pourquoi la dépression existe au niveau de l'espèce humaine mais donne du même coup un cadre pour comprendre pourquoi elle touche un individu. Alors que l'étude des mécanismes causaux de la dépression peut en expliquer des aspects variés dont certains sont strictement biologiques et la rapprochent d'une maladie non mentale mais physique ou somatique, cela laisse de côté les événements de la vie de la personne, mais aussi ses états psychologiques, ses intentions et ses traits de personnalité, qu'elle atteint pourtant spécifiquement en tant que trouble mental. Au contraire l'idée d'une tristesse pathologique appelle directement une recherche des déterminants individuels de cet état et de ses répercussions sur la vie de la personne.

- Troisièmement interpréter la dépression comme une tristesse pathologique, en permettant d'inscrire la dépression dans la vie mentale ou psychique du patient, plus particulièrement affective, réconcilie les deux grandes visions du trouble mental que sont celle d'un dysfonctionnement (défaillances de mécanismes) et celle d'un effondrement du sens (rupture de connexions significantes) (Bolton, 2008). La restauration de la cohérence de la vie psychique du patient atteint d'un épisode dépressif est au principe de la thérapie. Voir la dépression comme une émotion pathologique c'est la voir à la fois comme dysfonctionnement de cette vie psychique et comme un épisode de perte de sens, donc se donner les moyens d'agir à la fois au niveau des significations dont les liens font le cours de la vie mentale, et des mécanismes qui en sous-tendent la dynamique.

Limites de ce modèle évolutionniste

Malgré ces apports, un tel modèle évolutionniste de la dépression comme émotion pathologique me semble rencontrer deux types de limites : des limites de fait dans les explications qu'il donne de la dépression en faisant référence à la fonction évolutive des traits psychologiques et comportementaux qu'elle met en jeu, et une limite de principe dans son recours à l'évolutionnisme. Je dégage les limites de fait de ce modèle à partir de celles qui ont été mises en évidence par Luc Faucher pour d'autres modèles évolutionnistes de la dépression (Faucher, 2016) et je montre qu'il peut répondre à certaines, mais que ce n'est pas en faisant appel à l'évolutionnisme. En revanche le fait de reposer sur l'évolutionnisme, expose ce modèle à une « objection épistémique », également identifiée par Luc Faucher, et à laquelle il ne peut répondre qu'en ayant recours à la psychologie évolutionniste mais dans ce recours il rencontre une limite de principe (Faucher & Forest, 2021). Dans un dernier temps je regarderai si une théorie évolutionniste en philosophie des émotions permet de surmonter cette limite de principe.

a) Des limites de fait

Luc Faucher rappelle qu'une hypothèse évolutionniste à propos d'un trait doit non seulement pouvoir en expliquer toutes les caractéristiques par rapport à la fonction qui lui est hypothétiquement conférée, mais aussi être cohérente avec les connaissances issues des différents domaines scientifiques, dans le cas de la pathologie mentale : la psychologie, les neurosciences, l'éthologie, etc. Il dégage alors sept raisons pour lesquelles les modèles évolutionnistes de la dépression risquent de ne pas satisfaire ces contraintes et donc de ne pas être valides dans l'état actuel de la science à partir du modèle de Nesse (Nesse, 2009) et celui d'Andrews et Thompson (Andrews & Thomson Jr, 2009) qui n'interprètent pas la dépression comme une émotion. Voyons dans quelle mesure le modèle de Horwitz et Wakefield rencontre également ces limites.

1) D'après Luc Faucher, au vu de la diversité des routes étiologiques qui peuvent mener à la dépression, il serait improbable que l'unique hypothèse explicative issue d'un modèle évolutionniste soit à même de rendre compte de la dépression. La réponse de Nesse est que même si dans de nombreux cas la réponse dépressive n'est pas adaptée et se déclenche à cause d'une défaillance variable, cela n'empêche pas qu'il faille une explication évolutionniste à l'existence de cette réponse dépressive chez l'être humain et d'autres espèces, que son modèle fournit (Faucher, 2016). De même le modèle de Horwitz et Wakefield est compatible avec un très grand nombre de scénarios expliquant que la tristesse se produise de manière dysfonctionnelle, ne remplissant pas véritablement son rôle de réparation d'une perte.

2) Un modèle évolutionniste dit qu'une réponse normale, fonctionnelle, est proportionnée à son événement déclencheur, alors qu'une réponse disproportionnée, de ce fait inintelligible, est

dysfonctionnelle et peut être pathologique. Mais ce critère est peu convaincant selon Luc Faucher car l'estimation qu'une réponse dépressive n'est pas proportionnelle à ses causes reste subjective, et il y aura des cas où la réponse semblera proportionnée à ses causes mais sera tout de même pathologique : par exemple un deuil peut devenir un épisode dépressif. Luc Faucher cite une littérature selon laquelle une prise en compte de la phénoménologie affective serait mieux à même de départager états dépressifs pathologiques et tristesse normale. Comme cette critique vise directement la capacité d'un modèle évolutionniste à rendre compte du vécu affectif de la dépression comme pathologique, elle s'applique au modèle de Horwitz et Wakefield et il semble difficile d'y répondre. Que la tristesse ait pour fonction évolutive de répondre à une perte ne nous dit pas quand elle est proportionnée à la perte subie. Or si l'on caractérise la disproportion par les conséquences de la tristesse, par son retentissement pour le patient, sa définition par une fonction évolutive est superflue.

3) Les modèles évolutionnistes affirment permettre la prise en compte des événements de la vie dans le déclenchement de la dépression, alors que les neurosciences ne le permettent pas. Pour Luc Faucher cela ne rend pas compte des pratiques réelles des neurosciences qui intègrent tout de même des facteurs environnementaux dans leurs explications, susceptibles de faire une place à ces événements de vie. Mais seuls les modèles évolutionnistes permettraient en revanche de rendre compte de la dépression comme d'une réponse, susceptible d'être adaptative, à ces événements. Cette ligne argumentative est directement présente dans le modèle d'Horwitz et Wakefield. Alors ce modèle suppose que l'on puisse déterminer une fonction évolutive de la tristesse, et plus généralement des émotions.

4) Les modèles évolutionnistes examinés par Luc Faucher n'expliquent pas tous les symptômes de la dépression, donc ne remplissent pas l'exigence de rendre compte de toutes les caractéristiques du trait examiné par sa fonction adaptative. Cette critique s'applique au modèle de Horwitz et Wakefield dans la mesure où il y a dans la dépression des symptômes qui semblent effectivement difficiles à expliquer par une tristesse pathologique : dans certains tableaux cliniques de la dépression la tristesse est absente, et la dépression prend principalement la forme d'un « ralentissement psychomoteur » avec une baisse générale du comportement, de la motivation, de la prise de décision, de l'activité mentale (Widlöcher, 2002) ou bien le tableau clinique est marqué par une perte d'intérêt et de plaisir dominante plutôt que par une humeur triste dominante (Gaillard et al., 2013). Une réponse possible à cela depuis le modèle de Horwitz et Wakefield est que (i) la tristesse, tout en étant une adaptation, peut prendre des formes très variables car son phénotype s'explique par des facteurs biologiques mais aussi par des facteurs culturels qui ont introduit une variabilité dans ce trait au cours de la phylogenèse, (ii) il y a également une variabilité inter-individuelle de la manière dont les traits hérités de la sélection naturelle s'expriment, et enfin (iii) il y a des facteurs de variabilité au cours de l'ontogenèse.

5) Les modèles évolutionnistes ne sont pas entièrement cohérents avec les données scientifiques issues des sciences de base sur la dépression, ce que Luc Faucher affirme en citant des données qui contredisent directement l'idée que la dépression aurait la fonction évolutive supposée par le modèle d'Andrews et Thompson, celui de permettre l'analyse et la résolution de problèmes complexes. Par exemple il n'est pas évident qu'on puisse interpréter relativement à la thèse d'une réaction visant à la réparation d'une perte ce que l'on sait du rôle de la sérotonine (5-HT) dans la dépression, aussi bien grâce à l'efficacité des médicaments antidépresseurs, en particulier les inhibiteurs de recapture de la sérotonine (ISRS), que grâce à des modèles animaux, des observations en imagerie cérébrale et des études post-mortem chez l'être humain. Ce neurotransmetteur est impliqué dans de nombreuses fonctions du système nerveux central grâce à des récepteurs codés par de nombreux gènes, et aussi avec un rôle hormonal dans le système nerveux périphérique. Par

exemple il pourrait avoir un rôle à jouer dans l'hippocampe ; en conséquence, des déficits dans sa transmission, à cause de problèmes au niveau de ses récepteurs post-synaptiques (5-HT1A) résultant d'une expression défailante des gènes codant pour ces récepteurs, ou bien d'effets délétères du stress sur les structures de l'hippocampe, pourraient causer des dysfonctionnements de l'hippocampe directement responsables de phénotypes anxieux et dépressifs, voire du risque suicidaire, via une perturbation du rôle de l'hippocampe dans la communication entre régions limbiques et corticales du cerveau (Albert & Lemonde, 2004). Or les difficultés de communication entre régions limbiques et corticales du cerveau semblent plutôt associées à une altération de l'humeur et du comportement, qui entraîne des difficultés de régulation des émotions rendant compte d'une baisse générale des émotions, de la réactivité émotionnelle, caractéristique de la symptomatologie dépressive (Mayberg, 2007). Ces données vont donc à l'encontre d'une augmentation des émotions négatives comme la tristesse dans la dépression.

6) Les modèles évolutionnistes n'expliquent pas les comorbidités fortes de la dépression avec d'autres pathologies comme les troubles anxieux, en tout cas moins bien que les neurosciences d'après Luc Faucher. Au contraire un modèle évolutionniste de la dépression comme émotion pathologique ne peut-il pas expliquer des comorbidités liées au déroulement de la vie affective, ou à des traits de l'émotion pathologique ? Par exemple la comorbidité fréquente entre dépression et anxiété se comprend bien dans un modèle de la dépression comme tristesse pathologique, car dans la vie affective la tristesse peut susciter de l'appréhension et inversement, d'autant plus si l'une réagit à des situations de perte et l'autre à l'anticipation de ce type de situations. Si la dépression est une tristesse disproportionnée par rapport à une perte subie ou se produisant sans qu'une perte ait réellement été subie, elle peut constituer une réaction inintelligible pour autrui, et susciter de ce fait des comportements d'évitement à l'égard de la personne dépressive ; or la menace des liens sociaux de l'individu, le fait qu'il les perçoive comme rompus ou simplement mis en danger, déclenche une réponse de stress ou d'anxiété douloureuse (Panksepp, 2003). Mais dans un cas comme dans l'autre on voit que l'explication des comorbidités de la dépression, si elle repose en partie sur une définition de la tristesse, ne fait pas intervenir le cadre évolutionniste : le fait que cette définition soit une fonction adaptative de la tristesse ne joue aucun rôle dans le fait d'expliquer certains aspects de la dépression qui rendent compte de ses comorbidités.

7) Enfin Luc Faucher dit que les modèles évolutionnistes s'appuient sur le fait que certains traitements sont efficaces de manière cohérente avec leur hypothèse explicative sur la dépression, mais les stratégies thérapeutiques qu'ils favorisent ne sont pas nécessairement les plus efficaces, tandis que d'autres fonctionnent alors même que leur principe thérapeutique va directement à l'encontre de l'hypothèse explicative de ces modèles évolutionnistes.

Le modèle de la tristesse pathologique est compatible avec l'efficacité des thérapies psychanalytiques et psychodynamiques qui rétablissent l'état affectif du patient en lui permettant de découvrir les causes inconscientes ou incontrôlées de ses affects pathologiques (Desseilles, 2010), mais aussi avec l'efficacité de thérapies centrées sur la régulation des émotions comme les thérapies cognitivo-comportementales de la dernière vague dite « émotionnelle » qui travaillent sur la régulation des émotions. Les stratégies médicamenteuses peuvent être vues comme agissant sur un état neurobiologique corrélatif de la tristesse, causant ou résultant du fait qu'elle se déclenche sans répondre à sa fonction évolutive de réparation de la perte. Mais le fait que la définition de la tristesse soit sa fonction adaptative ne joue pas de rôle particulier dans le pouvoir heuristique du modèle de la tristesse pathologique par rapport à ces stratégies thérapeutiques, c'est seulement le fait qu'il mette une émotion et avec elle le vécu affectif au centre de la dépression, qui joue un rôle.

Il y a donc des limites de fait du modèle évolutionniste de la dépression comme tristesse pathologique auxquelles il lui semble difficile de répondre comme la difficulté à se concilier avec

certaines résultats des sciences de base sur la dépression. Il y a d'autres limites de fait auxquelles il semble pouvoir répondre mais sans que cela repose sur le cadre évolutionniste, comme le fait de rendre compte de certaines comorbidités de la dépression, ou de l'efficacité de certaines thérapies. Il y a enfin des limites de fait auxquelles il peut répondre si et seulement si le cadre évolutionniste est justifié, c'est-à-dire si on peut vraiment parler de la tristesse comme une adaptation, avec une fonction évolutive de réparation d'une perte qui rend compte de toutes ses caractéristiques en tant que trait phénotypique : le fait d'être une réponse adaptative, qui se concilie avec divers scénarios étiologiques de la dépression, et qui peut être interprétée par rapport à des événements de vie, le fait qu'il soit possible de définir ce qu'est une tristesse disproportionnée par rapport à sa cause pour tracer la limite du pathologique dans ce modèle, et enfin le fait de rendre compte de la variabilité des symptômes et tableaux cliniques de la dépression, par la variabilité de l'expression phénotypique de la tristesse.

Donc examiner les limites de fait du modèle évolutionniste de la dépression comme tristesse pathologique nous conduit à examiner une limite de principe, celle de la possibilité de faire appel à une fonction adaptative du trait par lequel il explique la dépression.

b) Une limite de principe

Une objection épistémique faite par Bergner (1997) pose problème pour une interprétation des troubles mentaux comme des dysfonctionnements par rapport à une notion évolutionniste de fonction. Cette objection est reprise par Luc Faucher (2021) qui montre que le modèle de Horwitz et Wakefield ne peut lui répondre qu'en s'appuyant sur la psychologie évolutionniste, or celle-ci ne rencontre une limite de principe montrée par Griffiths (1996) pour rendre compte de tous les traits psychologiques ou comportementaux affectés par les troubles mentaux.

L'objection de Bergner telle que la reprend Faucher (2021) est que l'analyse des troubles mentaux en termes de dysfonctionnement dépend de ce qu'on puisse mettre en évidence certains faits concernant des mécanismes internes qui sous-tendent notre fonctionnement psychologique normal, et reconstituer leur histoire évolutive. Or cela est parfois impossible. Faucher (2021) relève trois réponses de Wakefield à cette objection qui sont 1) que nous avons l'intuition de ce qui fonctionne et ne fonctionne pas, 2) que la fonction d'un mécanisme est parfois évidente et 3) que l'on a parfois besoin de la psychologie évolutionniste pour identifier la fonction de certaines composantes de notre esprit, voire pour identifier ces composantes. Mais Faucher (2021) rappelle que nos intuitions sur ce qui fonctionne ou ne fonctionne pas sont souvent bien trop réduites pour nous indiquer ce qui ne fonctionne pas, et ont souvent été trompeuses en matière de normalité psychologique, comme l'ont montré les nombreux cas de pathologisation abusive de traits normaux dans toute l'histoire de la santé mentale, par exemple la masturbation ou l'orgasme féminin. Deuxièmement il est rare que la fonction d'un trait soit évidente, surtout sa fonction adaptative, comme l'ont montré des erreurs historiques majeures d'interprétation de certains organismes. Il faut donc s'en remettre à une théorie qui détermine la fonction des mécanismes observés par leur histoire adaptative, ce qui concernant les traits ou mécanismes psychologiques, va être la psychologie évolutionniste (Faucher, 2021).

En effet la psychologie évolutionniste veut déterminer le fonctionnement normal de notre esprit en mettant en évidence ses mécanismes par rapport au rôle qu'ils ont historiquement rempli dans notre adaptation à notre environnement. Pour cela elle raisonne en termes d'adaptation, en isolant des problèmes posés par l'environnement et que l'organisme a dû devoir résoudre, et elle reconstitue l'histoire de nos traits psychologiques actuels d'après un scénario probable de la manière dont ils ont été sélectionnés dans un environnement d'origine puis conservés au cours de notre histoire évolutive. Donc la psychologie évolutionniste s'appuie sur une relation forte entre la forme biologique et les forces adaptatives : c'est cette relation qui justifie d'étudier l'architecture de l'esprit par rapport à des problèmes adaptatifs et de chercher des pressions expliquant les

mécanismes observés (Faucher, 2021). Mais une telle explication n'a de valeur objective que si l'on dispose d'un riche ensemble de connaissances historiques en particulier sur les conditions initiales ayant donné lieu aux adaptations : c'est l'argument de Griffiths (Griffiths, 1996) rappelé par Faucher. Les explications adaptatives d'une part s'appliquent à des traits ou des mécanismes concernant l'organisme et l'environnement dans lequel il a subi des évolutions, susceptibles de subir des variations sur lesquelles peut s'exercer la sélection naturelle, et d'autre part font intervenir un régime de sélection, des pressions adaptatives qui rendent compte de cette sélection. Dans une telle explication il faut donc appliquer une méthode que Griffiths qualifie d'historique et de comparative, consistant à rechercher les traits que possédait l'organisme quand les pressions sélectives se sont exercées sur lui, à comparer ces traits à ceux de ses ancêtres phylogénétiques proches, et à déterminer ces pressions sélectives par rapport à ce que l'on peut découvrir des conditions environnementales d'alors (Faucher, 2021).

La difficulté est alors qu'il n'est pas évident qu'il soit possible d'historiciser ainsi toute explication adaptative : Faucher (2021) s'appuie sur une littérature qui indique que pour un certain nombre de traits humains les données nécessaires pour appliquer la méthode comparative et historique décrite par Griffiths ne sont pas disponibles, voire resteront toujours hors de portée si elles résident dans un passé dont nous ne retrouverons jamais la trace. Pour ces traits il est donc impossible de déterminer leur fonction adaptative car nous n'aurons jamais accès aux faits qui permettraient vraiment d'établir qu'ils sont des adaptations, même si pour d'autres traits humains plus récents il semble en revanche possible d'établir un jour une explication évolutionniste historiquement prouvée. Peut-on alors trouver en philosophie des émotions une théorie prouvant que les émotions font partie des traits dont on peut donner une explication évolutionniste correcte ? Griffiths a lui-même appliqué son argument aux émotions : suivant l'argumentation de sa position « éliminativiste » à propos du concept d'émotion, toutes les émotions ne font pas partie des traits dont on peut avoir une explication évolutionniste convaincante suivant la méthode historique et comparative, mais seulement certaines émotions basiques qu'il appelle « programmes affectifs » (Griffiths, 1997). Dans la défense de sa théorie évolutionniste des émotions, Prinz (2004) a défendu, contre Griffiths, la possibilité de définir les émotions par une fonction adaptative. Je vais donc m'intéresser à la théorie évolutionniste des émotions de Prinz (2004) pour voir si elle peut ou non soutenir le modèle évolutionniste de la dépression comme tristesse dysfonctionnelle de Horwitz et Wakefield (Horwitz & Wakefield, 2007) en surmontant la limite de principe qui est qu'un tel modèle doit pouvoir rendre compte des émotions par leur fonction adaptative.

1. 3. La théorie évolutionniste des émotions de Prinz donne-t-elle un cadre à la définition de la tristesse pathologique relativement à une fonction adaptative ?

La théorie de Prinz (2004) offre à première vue un soutien possible au modèle d'Horwitz et Wakefield, puisqu'elle définit les émotions comme des adaptations. Les émotions sont des « évaluations incarnées » qui sont des perceptions d'états du corps indiquant un certain état des relations avec l'environnement, et qui ont été sélectionnées pour cela, elles sont donc définies par cette fonction adaptative.

a) Définition de la tristesse pathologique dans le cadre de la théorie évolutionniste des émotions de Prinz

Prinz reprend l'argument de la théorie James-Lange selon lequel les émotions font nécessairement intervenir une expérience de ressenti corporel, parce que si l'on élimine celle-ci par une expérience de pensée, il ne reste rien de l'émotion. Si cette théorie n'avait pas les moyens de montrer de façon convaincante que des réactions corporelles suffisamment distinctes existent pour donner lieu à toutes les émotions particulières distinctes, Prinz affirme cela est possible grâce aux

travaux de Damasio, en étendant les états du corps typiques des émotions à des états neuraux qui détectent ou simulent des changements corporels incluant des changements du milieu interne (Prinz, 2004 ; p. 58). De plus ces patterns d'activation étant corrélés à des situations typiques, ils signalent de manière fiable des relations entre l'organisme et son environnement, ce que Prinz présente comme le point de divergence entre sa théorie et celle de Damasio pour qui les émotions sont seulement des perceptions d'états du corps.

Les émotions sont alors des représentations de ces relations organisme-environnement au sens téléosémantique. Une représentation au sens téléosémantique, issu des travaux de Dretske et Millikan, satisfait deux conditions : contenir de l'information et pouvoir être juste ou erronée ; de plus une représentation dans un système biologique signale de manière fiable quelque chose dans l'environnement et a pour fonction adaptative de le signaler, c'est-à-dire qu'il s'agit d'un mécanisme biologique qui a été retenu par la sélection naturelle car il est utilisé par l'organisme comme indiquant de manière fiable quelque chose dans l'environnement (Chalmers, 2002). Prinz définit une représentation mentale comme un mécanisme cognitif de détection spécifique, plus précisément comme un état causé de manière fiable par quelque chose et mis en place par apprentissage ou bien au cours de l'évolution pour détecter cette chose (Prinz, 2004, p. 54). Les émotions sont alors des représentations en un sens téléosémantique parce qu'elles sont des perceptions d'états corporels qui véhiculent de manière fiable de l'information sur l'environnement, et ont été sélectionnées au cours de l'évolution pour cela. Cela fait reposer la définition des émotions comme des « évaluations incarnées » sur le cadre évolutionniste et c'est le fait même d'avoir été sélectionnées pour remplir cette fonction évolutive qui les définit comme des représentations évaluant l'état des relations entre l'organisme et l'environnement dans une situation donnée. De plus l'état corporel indiquant l'état de ces relations organisme-environnement est lié à des renforceurs positifs ou négatifs du comportement qui lui confèrent une valence : la représentation de la relation de l'organisme avec l'environnement est ainsi évaluée positivement ou négativement, d'où la notion d'« évaluation incarnée » qui résume la définition des émotions.

Prinz rend alors compte de tous les traits des émotions relativement à cette fonction évolutive.

- Les conditions de déclenchement des émotions ont l'air très variables, mais concernent toujours une relation avec l'environnement. Les relations organisme-environnement que les émotions représentent sont des schémas fondamentaux de relations avec l'environnement, notion empruntée aux travaux de Lazarus : Prinz prend en exemple la tristesse qui représente la perte (Prinz, 2004, p. 64).
- Les schémas relationnels sont indiqués par les états du corps propres aux émotions donc ne se reflètent pas directement dans leur structure interne : par exemple la tristesse représente une perte irréparable mais elle ne peut pas pour autant être décomposée en éléments qui seraient différentes parties de la représentation, les uns représentant la perte et les autres son côté irréparable (Prinz, 2004, p. 66)
- Pour rendre compte des émotions qui ne se déclenchent pas en réponse à des stimuli identifiables dans l'environnement immédiat, à des stimuli sensoriels directs, Prinz dit que des états cognitifs complexes, des images mentales voire des contenus conceptuels, peuvent devenir des conditions de déclenchement de certaines émotions par un apprentissage ou des modifications acquises au cours de l'évolution, par « calibration » (Prinz, 2004, p. 100)
- Que les émotions soit en partie déterminées par des normes culturelles, s'explique par le même type de processus d'acquisition ontogénétique ou phylogénétique qui font que l'environnement culturel détermine en partie les conditions de déclenchement des émotions ; par exemple le dégoût semble être universellement un état du corps qui représente quelque chose de physiquement ou moralement désagréable, mais les normes culturelles influencent la définition de ce qui est tel (Prinz, 2004, p. 140). La culture peut donc influencer le contenu réel des émotions en déterminant la calibration de certaines émotions par

l'association de l'état corporel typique à un certain type de cognitions, c'est-à-dire un ensemble de représentations mentales qui vont en devenir des conditions de déclenchement.

- La culture peut aussi influencer le contenu nominal des émotions, en déterminant en partie certains mécanismes corporels comme la respiration ou les expressions faciales. Prinz rend ainsi compte de la variabilité culturelle des émotions par des modulations dans l'expression d'un trait phénotypique panculturel : les « évaluations incarnées ». (Prinz, 2004, p. 142-150)
- Prinz fait enfin une typologie des états affectifs en différenciant notamment émotions et humeurs : à la suite de Lazarus il dit que les humeurs sont des évaluations de l'arrière-plan existentiel de nos vies, donc représentent le cours de notre vie en général, c'est-à-dire que même si elles peuvent être induites par des événements spécifiques, leur fonction est de nous renseigner sur notre situation dans le monde en général, et non sur une situation particulière. Les humeurs sont donc également des « évaluations incarnées », elles représentent une relation globale avec le monde donc détectent des changements globaux dans les relations de l'organisme avec l'environnement, alors que les émotions représentent des relations avec un objet ou un événement isolé du monde, détectant des changements localisés dans les relations avec l'environnement. Prinz prend à nouveau l'exemple de la tristesse, qu'il différencie de l'humeur dépressive : « La tristesse représente une perte particulière tandis que la dépression représente un combat perdu. » (Prinz, 2004, p. 185).

Le modèle de la dépression comme tristesse pathologique au sens d'une émotion ne remplissant pas sa fonction adaptative de permettre la réparation d'une perte (Horwitz & Wakefield, 2007) peut-il s'appuyer sur la théorie des émotions comme évaluations incarnées (Prinz, 2004) ? Oui dans la mesure où dans cette théorie les émotions sont définies par une fonction adaptative, à l'égard de laquelle elles peuvent être jugées dysfonctionnelles si elles cessent de la remplir. En particulier la tristesse est définie comme la représentation d'une perte au sens où elle est la perception d'un état du corps qui a été sélectionné pour indiquer à l'organisme ce type particulier de relation avec son environnement, donc elle peut être dysfonctionnelle si elle ne remplit pas cette fonction, par exemple si l'on est triste en l'absence d'une perte subie ou de manière disproportionnée par rapport à une perte subie. Une tristesse dysfonctionnelle pourra alors être pathologique si elle cause en outre un préjudice significatif à l'individu. La théorie des émotions comme évaluations incarnées soutient ainsi la définition d'une tristesse pathologique en donnant un cadre à la définition du dysfonctionnement de la tristesse qui est la base objective de son caractère pathologique, tristesse pathologique par laquelle Horwitz et Wakefield définissent la dépression.

La limite de principe du modèle de la tristesse pathologique était que sa validité est conditionnée par la possibilité de donner une explication évolutionniste du trait par lequel il cherche à expliquer la dépression, soit l'émotion de tristesse. La théorie évolutionniste des émotions comme « évaluations incarnées », qui les définit et les explique par une fonction adaptative, permet-elle à ce modèle évolutionniste de la dépression de surmonter cette limite de principe ? Puisque cette théorie des émotions les définit, en tant que traits psychologiques du phénotype humain, par une fonction adaptative, elle a un raisonnement conforme à la psychologie évolutionniste, donc elle est exposée à l'objection de Griffiths (Griffiths, 1996) contre celle-ci. Or Griffiths a lui-même appliqué cet argument aux émotions (Griffiths, 1997) et Prinz en tient compte dans la défense de sa théorie, dont on peut dégager une réponse à cet argument.

b) La théorie évolutionniste de Prinz parvient-elle à répondre à l'objection de Griffiths à une définition des émotions par la psychologie évolutionniste ?

L'objection de Griffiths repose sur le constat qu'au niveau explicatif auquel la psychologie évolutionniste veut définir les émotions, on ne dispose pas des connaissances scientifiques nécessaires pour aller au-delà d'une explication spéculative. Les niveaux explicatifs possibles pour

la psychologie, sont selon lui le « niveau écologique », où on explique les propriétés concernant les interactions entre les organismes et leur environnement, le niveau intermédiaire ou « niveau computationnel », où décrit la tâche accomplie par le système ou l'organisme qu'on étudie, et le « niveau de l'implémentation », celui du support physique des mécanismes accomplissant la tâche étudiée (Griffiths, 1997, Chapitre 8). Les différents concepts d'émotions peuvent désigner des catégories de phénomènes à des niveaux explicatifs différents : par exemple, le concept de « peur » peut désigner une catégorie ou espèce au niveau écologique qui serait « la réponse au danger », ou alors une catégorie propre aux humains et à leurs proches parents dans l'histoire de l'évolution, au niveau computationnel en termes de rôle fonctionnel dans la vie psychologique ou le comportement, ou au niveau de l'implémentation en termes de circuits neuraux (Griffiths, 1997, Chapitre 9). Or au niveau écologique, il n'existe pas de théorie scientifique des émotions, c'est-à-dire une théorie qui les définirait par une fonction adaptative, un rôle dans l'interaction entre l'organisme et l'environnement pour lequel elles auraient été retenues par la sélection naturelle. La conséquence en est que l'explication de telle ou telle émotion par telle ou telle fonction ne peut être que spéculative : aucune explication ne peut s'imposer mieux qu'une autre puisqu'aucune ne peut être confirmée ou infirmée et que toutes paraissent valables du moment qu'elles donnent un scénario crédible de la sélection naturelle de cette émotion (Griffiths, 1997, Chapitre 9). La seule manière d'obtenir une explication non pas spéculative mais factuelle serait d'avoir une approche comparative et historique qui combinerait la fonction adaptative supposée, c'est-à-dire l'hypothèse sur le rôle du trait qu'on cherche à expliquer concernant l'interaction entre l'organisme et l'environnement, avec différents types de données vérifiables (Griffiths, 1997, Chapitre 5). Griffiths applique donc ici aux émotions son argument concernant la nécessité d'une approche comparative et historique pour expliquer les traits humains comme des adaptations (Griffiths, 1996)

Il explique (Griffiths, 1997, Chapitre 5) qu'une explication évolutionniste doit être historique car les résultats des processus de l'évolution ne reflètent pas seulement les interactions entre un organisme et son environnement, mais aussi des facteurs stochastiques de sélection des traits au sein des populations, donc il faut utiliser des méthodes empiriques et comparatives pour les reconstituer. De plus les organismes vivants sont les descendants de lignées passées par différentes phases écologiques, donc il faut faire appel à la phylogénie pour restreindre les généralisations à propos des facteurs écologiques à des groupes taxonomiques avec un héritage commun. Cette approche comparative et historique a plusieurs avantages par rapport à la psychologie évolutionniste « spéculative ». Premièrement elle est plus méfiante envers la possibilité d'expliquer *a priori* des adaptations cognitives, puisqu'elle adopte plutôt une heuristique phylogénétique cherchant chez les êtres humains des systèmes cognitifs qui seraient des élaborations de mécanismes vus chez d'autres espèces et susceptibles de provenir d'un ancêtre commun. Deuxièmement elle fait une plus grande utilisation de données comparatives pour identifier les différents mécanismes psychologiques existant, plutôt que de le faire à partir de « contraintes de solvabilité », c'est-à-dire de généralités sur le développement de mécanismes pour répondre à des problèmes écologiques. Troisièmement elle utilise des méthodes empiriques et comparatives pour confirmer ou infirmer les scénarios adaptatifs construits à partir de régularités comportementales et de mécanismes cognitifs, plutôt que de s'en tenir à une réflexion sur leur caractère adaptatif censée révéler leur véritable but. Mais une telle explication adaptative historique des émotions fait encore défaut, et la psychologie évolutionniste se base uniquement sur l'écologie au sens de l'étude des relations entre l'organisme et l'environnement pour faire des hypothèses sur les fonctions adaptatives des émotions. La conséquence de cet argument de Griffiths est que toute définition des émotions conforme à la psychologie évolutionniste, soit par une fonction adaptative au niveau écologique, comme celle de Prinz, reste spéculative.

Griffiths ajoute que même si on parvenait à expliquer les émotions par la méthode comparative et historique, cela n'apporterait rien à notre compréhension des émotions aux deux autres niveaux explicatifs, c'est-à-dire sur leur rôle fonctionnel et leur implémentation neurale ou

neurobiologique (Griffiths, 1997, Chapitre 9). Or c'est aussi à ces niveaux explicatifs qu'une théorie des émotions est utile à un modèle évolutionniste de la dépression : les limites de fait que le modèle évolutionniste de la dépression ne parvenait pas à surmonter concernaient ces niveaux explicatifs puisqu'il s'agissait de s'accorder avec les résultats des sciences de base sur les mécanismes défailants dans la dépression, qui concernent par exemple la psychologie, donc le niveau fonctionnel, ou la neurobiologie, donc le niveau de l'implémentation. Mais définir les émotions par une fonction adaptative, c'est les définir comme une catégorie écologique, or une telle catégorie repose sur un ensemble de forces adaptatives qui façonnent le trait définissant cette catégorie, ce qui constitue un mécanisme trop éloigné du rôle fonctionnel et de l'implémentation physique de ce trait pour en rendre compte. Des principes écologiques généraux ne permettent pas d'expliquer les formes concrètes de l'expression des émotions, dont nombre d'éléments sont arbitraires, ce qui ferait en revanche l'objet d'une explication adaptative historique.

L'argumentation de Prinz pour défendre sa théorie offre une réponse à cet argument de Griffiths puisqu'elle défend à la fois la possibilité d'identifier une fonction adaptative des émotions et celle d'expliquer leurs traits par rapport à cette fonction (Prinz, 2004). Sa définition des émotions comme « évaluations incarnées » repose sur le cadre évolutionniste puisque les émotions sont des perceptions d'états du corps qui représentent des relations entre l'organisme et l'environnement, ayant pour fonction adaptative d'indiquer de manière fiable à l'organisme ces relations avec l'environnement en lui permettant d'y répondre. Prinz affirme qu'on peut dériver toutes les émotions d'un petit nombre d'émotions basiques caractérisées par des réactions corporelles dont il est clair qu'elles réagissent à un certain type de situation, donc représentent des schémas relationnels fondamentaux organisme-environnement. C'est de cet argument que dépend la possibilité de prouver scientifiquement que la fonction adaptative de toutes les émotions est de représenter les relations organisme-environnement par un état du corps, et qu'on peut rendre compte de leurs traits vis-à-vis de cette fonction. Mais comme aucun consensus scientifique n'a pu être atteint sur les émotions basiques depuis les travaux d'Ekman, Friesen et Izard, la liste des émotions basiques étant discutée tout autant que la définition même d'une émotion basique, il me semble que cet argument reste spéculatif.

Cela se confirme dans la mesure où Prinz présente lui-même la liste des émotions basiques, sur l'existence de laquelle sa théorie s'appuie, comme restant à établir. Selon lui la science devra combiner plusieurs approches, celle d'Ekman qui cherche des expressions faciales universelles, celle de Plutchik qui cherche des fonctions adaptatives fondamentales, celle de Panksepp qui cherche des circuits neuraux primitifs, ainsi que d'autres ensembles de données (données neurochimiques, données sur le développement, réflexion sur les concepts émotionnels, données introspectives sur les expériences émotionnelles, recherches linguistiques cross-culturelles...). Prinz propose lui-même une liste d'émotions basiques : la frustration, la panique, l'anxiété, le dégoût physique, la détresse de séparation, la conscience de soi aversive, pour les émotions basiques « négatives », et la satisfaction, la stimulation et l'attachement pour les émotions basiques « positives » (Prinz, 2004, p. 151-157). Il existerait un répertoire d'émotions basiques auxquelles ne correspondent pas nécessairement de concepts courants, et les émotions du répertoire humain adulte seraient à la fois biologiques et culturellement élaborées au cours de l'évolution de l'espèce et du développement de l'individu à partir de ces émotions basiques. Par exemple certaines émotions dériveraient d'émotions basiques en acquérant comme antécédents des contenus cognitifs spécifiques : la jalousie serait une réaction analogue à celle de la colère mais systématiquement associée au concept d'infidélité. Mais Prinz dit ici explicitement qu'il expose des spéculations appelées à être confirmées ou infirmées par des résultats scientifiques à venir (Prinz, 2004, p. 157)

Prinz dit qu'il « conjecture » que même si les avancées scientifiques dans ces différents domaines ne sont pas encore suffisantes pour permettre de faire une liste définitive d'émotions basiques, celles-ci seront dans tous les cas des « évaluations incarnées » (Prinz, 2004, p. 91). Mais

cela reste spéculatif car pour en donner la preuve, il faudrait disposer de cette liste, et que celle-ci rassemble des réactions corporelles indiquant chacune de manière fiable à l'organisme une certaine relation avec son environnement. De plus c'est l'ensemble de la théorie évolutionniste de Prinz qui est ainsi hypothéqué car pour les autres émotions du répertoire courant il faudrait ensuite pouvoir reconstituer l'histoire évolutive par laquelle elles ont dérivé des émotions basiques, c'est-à-dire appliquer à celles-ci la méthode comparative et historique préconisée par Griffiths, pour prouver qu'elles sont toutes des « évaluations incarnées », des représentations avec une valence, qui sont des perceptions d'états du corps indiquant de manière fiable un état des relations organisme-environnement car ayant pour fonction adaptative de le faire.

Cependant l'objection de Griffiths à la psychologie évolutionniste ne menace pas à ses yeux la possibilité de déterminer déjà dans l'état actuel des connaissances scientifiques une fonction adaptative de certaines émotions basiques, or celles-ci incluent la tristesse. Puisque la liste des émotions basiques est débattue, cela fait dépendre la définition évolutionniste de la tristesse par une fonction adaptative dont a besoin le modèle de la dépression comme tristesse pathologique de Horwitz et Wakefield d'un autre débat scientifique, non plus sur la psychologie évolutionniste, mais sur les émotions basiques.

c) Peut-on déterminer une fonction adaptative de la tristesse en tant qu'émotion « basique » ?

Griffiths défend l'existence de réactions basiques, pan-culturelles, qui interviennent typiquement dans certaines émotions, et dont on peut déjà reconstituer l'histoire adaptative : il les appelle « programmes affectifs » et les définit comme des réponses brèves et automatisées déclenchées par des perceptions à des stades précoces, et mises en œuvre dans des structures cérébrales anatomiquement anciennes que nous avons en commun avec d'autres vertébrés, et qui se mettent en place précocement dans le développement de l'enfant. Ces réactions sont typiques des émotions de la liste issue des travaux d'Ekman, Izard et Friesen : la surprise, la peur, la colère, le dégoût, la tristesse et la joie. (Griffiths, 1997, Chapitre 4)

Or selon Griffiths, ces réactions ont une fonction adaptative qu'on peut déjà identifier en reconstituant leur histoire évolutive. Les données existantes sur les bases neurales de ces réactions montrent qu'elles assurent le déclenchement de certains comportements essentiels de manière rapide et involontaire : ces données montrent par exemple que ces réactions viennent du « cerveau limbique », qui est la partie phylogénétiquement la plus ancienne du cortex, surmontant le tronc cérébral, avec un rôle crucial de l'hypothalamus et de l'amygdale, ce qui indique une histoire évolutive plus ancienne et un fonctionnement autonome par rapport à la régulation du comportement par des fonctions cognitives corticales. Chaque programme affectif est alors une réponse à des événements importants pour l'organisme au niveau écologique, ce qui rend plausible le fait qu'il ait pu être avantageux au cours de l'évolution que ces programmes évoluent en lien avec des mécanismes spécifiques d'interprétation de l'environnement local par l'organisme. (Griffiths, 1997, Chapitre 9). Cela constitue bien, pour les émotions basiques, un rôle au niveau écologique, soit une fonction adaptative.

Cela veut-il dire que la théorie des émotions de Griffiths (Griffiths, 1997) permet finalement de définir la tristesse par une fonction adaptative ? Il ne s'agirait pas d'une représentation de la perte par un état du corps comme dans le modèle de Prinz mais par exemple d'une réaction comportementale à la perte. Cela permettrait-il alors de définir un dysfonctionnement de la tristesse par rapport à cette fonction adaptative, comme le requiert le modèle évolutionniste de la dépression comme tristesse pathologique de Horwitz et Wakefield ? Ce n'est pas sûr dans la mesure où ce qui fait l'objet d'une explication évolutionniste c'est une réaction typique de la tristesse mais qu'il me semble difficile de considérer comme équivalente à la tristesse elle-même, comme définition de celle-ci. En d'autres termes, la tristesse fait typiquement intervenir une certaine réaction basique dont on peut reconstituer l'histoire évolutive, mais celle-ci n'est qu'un des traits de la tristesse elle-

même, comme les expressions faciales, donc la fonction adaptative de cette réaction n'est pas celle de la tristesse. En effet on peut expliquer par l'évolution l'existence de certains traits intervenant dans les émotions, notamment dans leur expression, mais cela n'est jamais une explication des émotions elles-mêmes. Darwin avait mis en évidence trois principes de la sélection naturelle de l'expression des émotions chez l'animal et chez l'homme (Darwin, 1877) : (i) le principe des habitudes selon lequel des mouvements utiles à la communication interindividuelle ont été sélectionnés et sont devenus réflexe, comme les mimiques faciales, (ii) le principe de l'antithèse selon une réaction de l'organisme à une situation déclenche une réaction opposée lorsque la situation initiale cesse, et (iii) le principe d'action directe selon lequel l'activité du système nerveux se traduit directement par certaines manifestations comme le rougissement, l'augmentation du rythme cardiaque, etc. Ces trois principes expliquent bien les mécanismes de sélection de certaines manifestations émotionnelles dans le comportement humain, mais non l'émotion elle-même par une fonction évolutive. Ainsi les travaux d'Ekman, Izard et Friesen prolongent ceux de Darwin pour donner l'explication évolutionniste complète de certains traits comme certaines expressions faciales, mais c'est la fonction adaptative de cette expression seulement qui est donnée. De tels travaux ne suffisent pas à définir une émotion au niveau écologique, par une fonction adaptative globale. Ainsi l'unité même de chaque émotion basique, lui permettant de constituer un trait biologique hérité de la sélection naturelle, ne fait pas consensus aujourd'hui, ni *a fortiori* la liste de ces émotions basiques (e.g. Barrett, 2006; Colombetti, 2009)

Griffiths dit que le contenu intentionnel d'une émotion suivant son concept dans la langue courante peut renvoyer au fonctionnement de deux systèmes distincts que sont une réaction basique modulaire, dont on connaît l'histoire évolutive, ou bien des processus cognitifs de haut niveau, dont on ignore l'histoire et donc la fonction adaptative. Ces deux systèmes peuvent s'accorder ou bien être en désaccord et dans ce cas, si le premier commande une réponse émotionnelle alors que les processus cognitifs de haut niveau évaluent ses stimuli de façon contraire, on aura une émotion réflexe, « irrationnelle », tandis que si les processus cognitifs de haut niveau procèdent à une évaluation alors que le mécanisme modulaire n'a pas été activé, on aura une évaluation sans émotion ; enfin les émotions sans objet correspondraient à des activations inappropriées du mécanisme modulaire (Griffiths, 1997, p. 98-99). La tristesse impliquerait donc tout autant une réaction basique héritée de la sélection naturelle, dont l'antécédent typique semble être une situation de perte, qu'une évaluation par des représentations à un haut niveau cognitif, qui peuvent être formellement liées à la perte mais dont on ignore l'histoire évolutive. *Mais alors le cadre évolutionniste est superflu, puisqu'on peut définir finalement la tristesse comme l'émotion qui répond à une perte, sans qu'on désigne par là une fonction adaptative, même si c'est peut-être celle d'une réaction comportementale basique qui intervient typiquement dans la tristesse.*

Ne peut-on dès lors envisager un modèle de la dépression comme tristesse dysfonctionnelle sans arrière-plan évolutionniste, simplement comme une tristesse ne réagissant pas, ou de façon disproportionnée, à une perte ? Mais si le fait de représenter la perte n'est pas la fonction adaptative de la tristesse, parler ici d'un dysfonctionnement suppose une théorie de la vie mentale ou affective, dans laquelle la fonction de la tristesse est de répondre à la perte, non en un sens évolutionniste de cette fonction mais au sens d'un rôle fonctionnel, c'est-à-dire d'une contribution au fonctionnement de cette vie mentale ou affective. Préciser ce rôle fonctionnel, cette contribution, consiste alors à spécifier l'ensemble des causes et des effets de la tristesse, ou plus généralement des émotions dans la vie mentale. Cela va supposer une théorie des émotions qui les analyse en composantes articulées les unes aux autres et aux autres états mentaux pour reconstituer les mécanismes par lesquels elles sont causées et ont des effets, mécanismes dont la défaillance constitue un dysfonctionnement. C'est par ce type de théorie que la philosophie des émotions me semble pouvoir apporter quelque chose à l'étude de la dépression sous l'angle de la médecine du « comment ».

2. Quel cadre heuristique la philosophie des émotions peut-elle offrir à un modèle non évolutionniste de la dépression comme trouble affectif ?

Si le cadre évolutionniste est superflu ou spéculatif pour définir les émotions, il reste possible d'évaluer leur contribution à la vie psychologique non en termes de fonction adaptative mais en termes de rôle fonctionnel, c'est-à-dire des relations qu'elles entretiennent avec les autres états affectifs, les états mentaux, les stimuli, le comportement. Cela entre dans la définition d'une fonction comme contribution d'un composant à l'explication d'une certaine capacité d'un système (Cummins, 1975), définition applicable à une notion de fonction et de dysfonction adaptée pour définir des états de santé et de pathologie ou de maladie (Forest & Le Bidan, 2016). Cela laisse envisager la possibilité d'identifier dans la dépression des dysfonctionnements de la vie affective susceptibles de sous-tendre sa définition comme trouble mental. L'approche du trouble mental comme dysfonctionnement de la vie psychologique au sens d'une défaillance des relations entre états mentaux, est un paradigme majeur de la psychiatrie contemporaine, supposant la possibilité d'identifier un fonctionnement psychologique « normal » pour en distinguer des états dysfonctionnels pathologiques (Bolton, 2008). Si certains états affectifs s'inscrivent dans des dysfonctionnements de la vie psychologique qui sous-tendent des états pathologiques, repérer ces états dysfonctionnels suppose une caractérisation précise de cette vie psychologique (Forest, 2016) donc d'étudier les mécanismes affectifs propres à la dépression, puis de s'appuyer sur une théorie des états affectifs permettant de caractériser ce qui dysfonctionne dans ces mécanismes. La philosophie des émotions propose de telles théories donc peut contribuer à la compréhension de la dépression. Une telle contribution ne peut avoir le caractère intégrateur d'un modèle évolutionniste de la dépression comme émotion pathologique, mais une grande partie des recherches scientifiques étiologiques sur la dépression porte également sur les causes non pas lointaines mais proches de celle-ci, donc sur des mécanismes défaillants. Je vais d'abord situer les émotions dans ces recherches étiologiques. Puis je vais dégager un modèle fonctionnel de la dépression qui caractérise les mécanismes affectifs en termes de relations entre stimuli, états mentaux et comportements. Enfin je vais chercher en philosophie des émotions un cadre heuristique pour penser le vécu affectif de la dépression comme un dysfonctionnement de la vie affective.

2. 1. La place des émotions dans l'étude des causes de la dépression

L'orientation des recherches « étiopathogéniques » sur la dépression n'est pas de se demander « pourquoi » mais bien « comment » la dépression arrive, et de chercher des mécanismes défaillants derrière ses différents symptômes en vue d'améliorer diagnostic, pronostic et thérapeutique. Quel est alors la place des émotions dans la symptomatologie de ce trouble mental qui produit une très grande variété de tableaux cliniques ?

Les sciences cliniques, qui étudient les critères de définition et de différenciation des tableaux cliniques de la dépression et servent de point de départ aux recherches étiologiques, opérationnalisent la symptomatologie dépressive par des outils dont ressortent trois pôles : un pôle somatique, un pôle cognitif et comportemental, et un pôle affectif. Premièrement les sciences cliniques opérationnalisent la symptomatologie dépressive par des outils quantitatifs, principalement les échelles de mesure de la dépression dont ressortent les différents pôles de symptômes. Les principales échelles sont la *Hamilton Rating Scale for Depression* et la *Beck Depression Inventory*. La première, développée de 1960 à 1980, sous la forme d'un questionnaire avec choix de réponse permettant d'établir un score final, mesure la sévérité de l'épisode dépressif à partir de celle de ses symptômes : altération de l'humeur, anxiété, troubles du sommeil, perte de poids (Hamilton, 1986; Miller et al., 1985; Sharp, 2015). La seconde est également un questionnaire développé en 1961 par Aaron Beck dont les travaux ont inauguré une approche cognitive de la

dépression, et évalue également la sévérité de celle-ci en évaluant notamment deux facteurs qui sont la vision négative du soi (culpabilité, dévalorisation) et le fonctionnement somatique et physique (perte d'appétit, fatigue, insomnie...) (Beck et al., 1961, 1988). L'une et l'autre peuvent être complétées par d'autres questionnaires plus récents par exemple la *Center for Epidemiologic Studies Depression Scale* qui évalue plusieurs facteurs : les émotions négatives (affect négatif), l'absence d'émotion positive (anhédonie c'est-à-dire perte d'intérêt et de plaisir) et les symptômes somatiques (Carleton et al., 2013). La symptomatologie affective est donc à la fois au centre de la dépression et prise en compte sous l'angle de ses rapports avec le somatique, le cognitif et le comportemental. Deuxièmement les sciences cliniques partent de cette symptomatologie opérationnalisée, pour caractériser la dépression dans la population, par exemple étudier ses comorbidités avec d'autres troubles comme l'anxiété (Brown et al., 2001).

C'est donc la symptomatologie qui est le point de départ des sciences cliniques, plus particulièrement des recherches étiopathogéniques, qui partent d'hypothèses faites grâce à la caractérisation épidémiologique et clinique de la dépression pour en chercher un modèle étiologique : par exemple des différences cliniques entre la dépression chez les hommes et chez les femmes motivent la recherche de deux modèles étiologiques développementaux distincts (Kendler et al., 2002, 2006). Or l'organisation de ces recherches étiopathogéniques se fait suivant les différents pôles de la symptomatologie : par exemple sur le pôle somatique, des études portent sur le dérèglement du rythme circadien dans la dépression (Azorin & Kaladjian, 2009), sur le pôle comportemental, des recherches concernent sur les mécanismes psychologiques et biologiques de la production de l'effort (Cléry-Melin et al., 2011), sur le pôle cognitif des travaux sont faits sur différentes atteintes des fonctions exécutives comme l'attention ou la prise de décision (Polosan et al., 2016), et sur le pôle affectif des recherches visent les problèmes de régulation des émotions (Mirabel-Sarron, 2012) Ces recherches se font de plus aux différents niveaux explicatifs de la psychiatrie : le niveau comportemental que le développement, le niveau moléculaire, par exemple sur le rôle de la sérotonine (Albert & Lemonde, 2004), le niveau physiologique, par exemple sur les défaillances du système immunitaire (Guilbaud et al., 2002), le niveau génétique, par exemple sur les facteurs environnementaux de l'expression de certains gènes (Lesch, 2004), le niveau neural, notamment sur la neurobiologie de l'humeur (Mayberg, 2007)... L'horizon de ces recherches étiopathogéniques est d'intégrer ces différents niveaux explicatifs dans un modèle pluraliste de la dépression, intégration que doivent permettre de nouveaux types de classification des troubles mentaux comme le projet RdoC (Faucher & Goyer, 2015).

L'unification des niveaux explicatifs doit permettre de rendre compte de ce que l'on observe dans la clinique, donc du déroulement d'un épisode dépressif qui est un épisode thymique, un épisode d'altération de l'humeur, avec plus précisément une humeur dépressive qui est une humeur basse. La psychopathologie décrit celle-ci sur le plan psychologique, ou mental, par différentes dimensions que sont la perte d'intérêt et de plaisir ou anhédonie, l'émoussement affectif qui est une baisse générale voire une absence de réactions affectives, et l'alexithymie qui est l'absence de verbalisation ou plus généralement de manifestation symbolique des états affectifs (Pasquier, 2012). La physiopathologie cherche au contraire à caractériser et expliquer la symptomatologie somatique de la dépression. Je me concentre sur le plan mental plutôt que somatique, puisque j'aborde la dépression comme un trouble mental afin d'envisager des contributions de la philosophie des émotions à sa compréhension. Je laisse donc de côté la physiopathologie.

En psychopathologie de la dépression, des modèles fonctionnels expliquent le vécu affectif, caractérisé par les dimensions citées, par des relations entre états affectifs, cognitions et comportement : ce sont les modèles cognitifs. Comme ils décrivent des mécanismes, ils ne peuvent en caractériser des défaillances qu'en s'appuyant implicitement en arrière-plan sur une théorie du fonctionnement normal de la vie psychologique, qui soit une théorie du rôle fonctionnel des états affectifs.

Je vais dégager de la littérature un modèle cognitif consensuel de la dépression, puis chercher dans les théories de la philosophie des émotions qui ont aussi une approche fonctionnelle un cadre heuristique pour comprendre ce modèle cognitif comme un modèle du vécu affectif de la dépression en tant que caractérisé par des défaillances à l'égard d'un vécu psychologique normal.

2. 2. Un modèle fonctionnel cognitif de la dépression

On peut dégager des travaux récents sur cette symptomatologie dépressive un modèle fonctionnel cognitif qui met en évidence des interactions entre l'humeur dépressive, les émotions, les cognitions et le comportement qui en résulte.

L'humeur est « un état affectif qui bouge lentement sur un ou deux jours, et faiblement lié à des objets ou situations spécifiques » et les émotions sont « des réactions adaptatives, rapides, qui interviennent en réponse à de multiples stimuli spécifiques », dont l'expression est à trois niveaux « subjective, physique (tachycardie, sueurs...) et comportementale. » (Mirabel-Sarron, 2012). Le vécu dépressif est caractérisé par l'humeur dépressive qui est une humeur triste et par une diversité d'émotions négatives comme la colère ou l'irritabilité, la honte, la culpabilité, et formellement par l'érousement affectif, la perte d'intérêt et de plaisir ou anhédonie, et l'alexithymie. Un consensus se dégage autour du modèle dit de l'insensibilité émotionnelle au contexte pour rendre compte de ce vécu dépressif, décrivant une réactivité réduite à la fois aux stimuli positifs et négatifs, par rapport à deux modèles concurrents qui sont l'atténuation positive (réactivité réduite aux stimuli positifs), et la potentialisation négative (réactivité accrue aux stimuli négatifs) (Rottenberg et al., 2005). Or cette insensibilité émotionnelle au contexte peut être mise en relation avec des défaillances de régulation des émotions, via des déficits cognitifs, causés par l'humeur dépressive qu'ils entretiennent en retour.

a) L'insensibilité émotionnelle au contexte

Suivant cette thèse la dépression majeure est associée avec une réactivité émotionnelle réduite, y compris aux contextes tristes. C'est contre-intuitif puisque l'humeur dépressive est vue comme une humeur triste donc on pourrait penser qu'elle accroît la sensibilité aux stimuli qui lui sont congruents, c'est-à-dire suscitant des émotions négatives notamment la tristesse. Mais cette thèse est mieux soutenue que les deux thèses concurrentes : celle de l'atténuation des émotions positives et celle de la potentialisation négative (Rottenberg et al., 2005) par les données empiriques sur la façon dont les personnes dépressives réagissent aux stimuli tristes par rapport aux sujets contrôle. Il y a effectivement une atténuation des émotions positives puisque les personnes dépressives présentent une réactivité émotionnelle atténuée aux stimuli émotionnels positifs, ainsi que des symptômes de diminution de la motivation, qui peuvent être interprétés comme une tendance réduite à chercher à obtenir des choses positives ou des récompenses, mais cela est également en accord avec la thèse de l'insensibilité émotionnelle au contexte, alors qu'inversement la thèse de l'atténuation des émotions positives ne rend pas compte de la moindre réactivité des personnes dépressives aux stimuli négatifs (Rottenberg et al., 2005). L'hypothèse de la potentialisation négative prévoit que les personnes dépressives vont présenter une réactivité émotionnelle plus importante aux stimuli émotionnels négatifs. Cette hypothèse présente dans les premiers modèles cognitifs de la dépression issus des travaux de Beck (Beck, 1967) décrit un renforcement mutuel de l'humeur dépressive et des émotions négatives via l'activation de structures cognitives qui favorisent la prise en compte des stimuli négatifs et des émotions typiques de la dépression (notamment la tristesse). Mais cette hypothèse n'a finalement pas été très soutenue empiriquement : des biais en faveur des stimuli négatifs ne se traduisent pas nécessairement par davantage d'émotions négatives en réponse à ces stimuli (Rottenberg et al., 2005).

Les données expérimentales montrent donc que les personnes dépressives manifestent moins d'émotions car elles réagissent moins à un certain nombre de stimuli qui normalement provoquent

des émotions, si bien que la dépression se caractérise plutôt par une réactivité émotionnelle réduite, y compris aux contextes tristes (Rottenberg, 2005). Cette altération de la réactivité émotionnelle est de plus corrélée à la sévérité de la dépression : les personnes qui traversent un épisode de trouble dépressif majeur, indépendamment de la sévérité et de la durée de cet épisode, auront un résultat clinique meilleur s'ils conservent des réactions émotionnelles plus marquées aux événements, tandis qu'une absence relative de réponse à l'environnement peut contribuer au caractère chronique du trouble dépressif majeur (Peeters et al., 2010).

b) L'insensibilité émotionnelle au contexte et la régulation des émotions

L'« insensibilité émotionnelle au contexte » se révèle résulter de dysfonctionnements de la régulation des émotions. Cette dernière désigne la mise en œuvre de stratégies permettant de moduler l'émotion, concernant son occurrence, sa nature, son intensité et sa durée (Baeyens, 2019). Une importante littérature montre que les difficultés de régulation des émotions sont un élément prédictif de la dépression. On peut distinguer des stratégies « adaptatives » de régulation des émotions (résolution de problèmes, réévaluation et acceptation) et des stratégies « défaillantes » (suppression de l'émotion ou de la pensée, évitement de l'expérience ou du comportement, et rumination) qui sont des facteurs de vulnérabilité aux troubles affectifs (Aldao et al., 2010). Par exemple la tendance à recourir à la suppression semble favoriser la dépression (Ehring et al., 2010), ou bien plus généralement la tendance à recourir à des stratégies générées avant que l'émotion n'émerge (suppression, évitement), par opposition à des stratégies de contrôle du déroulement de l'émotion (Gross & John, 2003). Globalement les personnes dépressives se caractérisent sur le plan affectif par des des stratégies qui diminuent leur réactivité émotionnelle donc augmentent leur insensibilité émotionnelle au contexte, mais aussi leurs capacités de régulation des émotions négatives, dont le maintien entraîne à terme le maintien de l'humeur dépressive qui entraîne à son tour l'insensibilité émotionnelle au contexte. Par exemple une étude répertorie neuf stratégies cognitives de « coping » différentes que les personnes peuvent mettre en œuvre face à un événement négatif (l'auto-accusation, l'accusation d'autrui, la rumination, le catastrophisme, l'acceptation, la mise en perspective, la re-focalisation positive, le recentrage sur la planification, et la réévaluation positive) : l'importance de la rumination, du catastrophisme et un déficit de réévaluation positive prédisent de manière fiable la prédominance des émotions négatives, or ces stratégies sont plus fortement présentes chez les personnes dépressives (Martin & Dahlen, 2005). La rumination est une réponse aux émotions négatives qui gêne particulièrement la réparation de l'humeur et tend à prolonger les épisodes dépressifs ; elle est définie par un certain style de pensée plutôt qu'un contenu négatif, à savoir un processus de pensées récurrentes et d'idées qui tendent à se « recycler », c'est-à-dire à se maintenir et persister (Joormann & Gotlib, 2010). En définitive, les personnes dépressives et non dépressives pourraient donc ne pas tant différer par leur réponse initiale à un événement négatif, que par leur capacité à restaurer leur humeur après avoir fait l'expérience de la tristesse ou d'autres émotions négatives (Joormann, 2010). Ces défauts de régulation affective s'expliquent alors par des déficits cognitifs (Baeyens, 2019).

c) Les mécanismes cognitifs de la régulation affective

Les difficultés de régulation des émotions par lesquelles se traduit l'insensibilité émotionnelle au contexte caractéristique de la dépression proviennent de mécanismes cognitifs défaillants. Or ces troubles cognitifs sont causés par l'humeur dépressive, symptôme central de la dépression, ou inversement les défaillances cognitives peuvent être premières et entraîner en premier lieu l'installation de l'humeur dépressive, via des émotions négatives non régulées qui la causent. Quels sont ces mécanismes cognitifs intervenant dans la régulation des émotions ?

La présence de l'humeur dépressive s'accompagne de perturbations de fonction cognitives notamment l'attention, la mémoire et les fonctions exécutives qu'on peut répartir en inhibition, mise à jour, flexibilité cognitive (mémoire de travail), planification, ou encore fluence verbale (Polosan

et al., 2016). Par exemple l'humeur dépressive cause un biais attentionnel envers les indices émotionnels négatifs, un biais attentionnel symétrique de négligence des indices émotionnels positifs, et une mémorisation accrue des matériaux émotionnels négatifs (Leppänen, 2006). L'humeur dépressive cause également des biais mnésiques, notamment un biais favorisant la mémorisation de matériaux congruents à l'humeur dépressive donc négatifs, qui apparaîtrait de manière automatique ; dans des dépressions peu sévères, l'absence de biais mnésique semble résider dans le fait que les patients sont capables de développer des stratégies défensives conscientes contre l'effet de congruence avec l'humeur dépressive qui se joue au niveau de mécanismes mnésiques automatiques ; dans des dépressions plus sévères, les patients ne parviennent pas à mettre en œuvre des stratégies pour neutraliser ce biais donc pour évacuer le matériel négatif de leur mémoire (Colombel, 2007). L'humeur dépressive atteint aussi les fonctions exécutives, notamment l'inhibition, qui fait partie des processus de contrôle exécutif. Ces processus sélectionnent et mettent à jour le contenu de la mémoire de travail, système à capacité limitée déterminant l'orientation attentionnelle et l'activation temporaire de représentations conscientes (Joormann, 2010; Joormann & Gotlib, 2010). Un déficit d'inhibition est alors responsable du maintien d'un matériau négatif dans la mémoire de travail. Cela affaiblit les capacités de régulation des émotions en augmentant les tendances à la rumination et contribue donc à la persistance voire à l'intensification des émotions négatives (Joormann, 2010; Joormann & Gotlib, 2010). Le maintien de contenus négatifs en mémoire de travail augmente aussi la mémorisation à long terme de ce type de contenus, et gêne l'accès à des contenus positifs, par le rappel de souvenirs positifs par exemple, ce qui tend à maintenir l'humeur dépressive en empêchant le recours à des stratégies de régulation permettant de changer d'humeur ou d'émotions. La rumination, stratégie défailante de régulation des émotions, est ainsi favorisée par les déficits d'inhibition qui entraînent le maintien d'informations négatives en mémoire de travail, alors que des stratégies plus adaptées comme la réévaluation ou la distraction par exemple supposent que ces contenus négatifs soient inhibés (Baeyens, 2019; Joormann, 2010).

En amont de ces déficits cognitifs causés par l'humeur dépressive, des déficits cognitifs peuvent perturber directement la régulation des émotions, entraînant une humeur dépressive qui causera à son tour d'autres déficits cognitifs. Un bon exemple est donné par les anomalies de traitement des expressions faciales. On peut montrer des difficultés des sujets dépressifs à distinguer des expressions faciales, plus précisément une incapacité à identifier de manière adéquate des changements légers d'expression faciales, et un biais défavorable à l'identification comme « joyeuses » d'expressions faciales modérément joyeuses (Surguladze et al., 2004). Des expressions faciales positives (joyeuses), neutres ou ambiguës ont tendance à être évaluées davantage comme tristes et moins comme joyeuses par les sujets dépressifs, qui montrent aussi une plus grande vigilance et une sélectivité de l'attention en direction d'expressions tristes par opposition à des expressions joyeuses (Bourke et al., 2010). Les sujets dépressifs identifient les visages tristes et en colère aussi vite que les sujets contrôle mais mettent plus de temps à identifier les visages joyeux, ce qui soutient la thèse d'un défaut d'identification des émotions positives ; ou bien cela pourrait refléter un déficit non pas de détection des stimuli émotionnels positifs mais de mise en œuvre de la réponse émotionnelle à ces stimuli, cette réponse pouvant par exemple être retardée dans l'attente d'une confirmation plus importante de ces stimuli par les sujets dépressifs par rapport aux sujets non dépressifs (Joormann & Gotlib, 2006).

L'étude des mécanismes reliant cognitions, émotions et humeur dépressive constitue un modèle fonctionnel cognitif de la dépression. Celui-ci en décrit les mécanismes typiques chez des personnes diagnostiquées ou identifiées comme dépressives relativement à des sujets contrôle. Un tel modèle s'accorde bien avec une théorie fonctionnelle des états affectifs, qui définit l'humeur dépressive et les émotions comme des états fonctionnels, c'est-à-dire en termes de relations entre des stimuli, des états mentaux et des comportements.

d) Du modèle cognitif de la dépression à une théorie fonctionnelle de l'humeur

Du modèle cognitif il ressort que les sujets dépressifs ont des biais cognitifs en faveur des stimuli émotionnels négatifs et ont recours à des stratégies défaillantes de régulation des émotions comme l'évitement ou la suppression, qui rendent compte de la baisse de leur réactivité émotionnelle, ou les ruminations qui diminuent d'autant plus leurs ressources cognitives pour réguler leurs émotions. Ces stratégies défaillantes sont mises en œuvre à cause de déficits cognitifs, par exemple au niveau de fonctions exécutives comme l'inhibition. Les déficits cognitifs sont causés par la présence de l'humeur dépressive tout autant qu'ils la causent en retour : cette humeur apparaît comme l'origine de ces déficits cognitifs mais aussi comme la conséquence des émotions négatives mal régulées et des pensées négatives persistantes à cause de ces déficits, tant les biais attentionnels ou mnésiques que les déficits de fonctions exécutives. Le cercle vicieux entre humeur dépressive, émotions négatives et déficits cognitifs se traduit donc par une circularité du modèle fonctionnel cognitif du vécu dépressif.

Les théories fonctionnelles de l'humeur et des émotions qui les définissent et les distinguent par des rôle fonctionnels, donc en termes de causes et d'effets parmi les stimuli, les états mentaux et le comportement, définissent l'humeur non pas comme un état fonctionnel de premier ordre par des relations causales avec d'autres états mentaux, des stimuli et des comportements, mais comme un état fonctionnel global, c'est-à-dire un état des relations causales entre l'ensemble des stimuli, des états mentaux et du comportement. L'humeur n'est donc pas sur le même plan que les émotions, elle est un état d'ordre supérieur, ou de second ordre, ou dispositionnel (Siemer, 2009). Dans une telle théorie, la circularité des relations entre humeurs, émotions, et cognitions dans le modèle cognitif est apparente, et reflète le fait qu'en réalité l'humeur *consiste dans* les relations entre émotions et cognitions elles-mêmes. C'est *par définition* que l'humeur est là aussi longtemps que le cercle vicieux entre émotions négatives et cognitions négatives est présent, puisqu'elle est justement définie par l'état des relations entre les unes et les autres.

La principale théorie est celle de Griffiths qui définit l'humeur comme un état fonctionnel global qui dévie de l'état fonctionnel global normal de l'individu (Griffiths, 1997). Cela a trois conséquences : 1) il existe une description fonctionnelle de l'individu qui est son état fonctionnel global normal, et des descriptions fonctionnelles déviantes, correspondant aux différentes humeurs (notamment humeur dépressive, anxieuse ou euphorique) ; 2) une théorie des humeurs relie des inputs divers (environnementaux, etc.) avec des outputs qui sont le fait que la personne actualise une certaine description fonctionnelle globale ; 3) l'humeur est un état « d'ordre supérieur » par rapport aux émotions, c'est-à-dire que sa description fonctionnelle peut englober comme parties constitutives des émotions (elles-mêmes définies comme des états fonctionnels). Cette conception est à la fois proche du sens du concept courant d'humeur, et compatible avec ce que l'on sait de l'implémentation neurale des humeurs, puisque les humeurs semblent correspondre à des états chimiques du système nerveux central qui affectent la probabilité que certaines transitions entre certains états neuraux se produisent (Griffiths, 1997). Définir l'humeur comme un état fonctionnel d'ordre supérieur signifie donc qu'elle consiste dans des processus : en attendant d'identifier ces derniers sur le plan neurobiologique, elle est définie par un ensemble de changements concernant des processus cognitifs comme l'attention, la mémoire, la formation de concepts (Sizer, 2000).

Cette théorie est parfaitement adaptée à la conception de l'humeur qui se dégage des sciences cliniques, car Griffiths veut rendre compte de trois humeurs prototypiques qui sont l'euphorie, l'anxiété et la dépression (Griffiths, 1997). Mais cette approche fonctionnelle ne suffit pas à une définition de la dépression par un dysfonctionnement de la vie affective, puisqu'elle permet de décrire abstraitement ce qu'on observe dans un modèle fonctionnel comme le modèle cognitif de la dépression, à savoir la présence d'une humeur qui est un certain état des relations entre émotions et cognitions, mais non de déterminer si cela est dysfonctionnel et si oui de quelle nature est le dysfonctionnement. Pour cela, il nous faut une définition substantielle des relations entre émotions, humeur et cognitions dans la vie psychologique « normale ».

Une telle définition est donnée par les théories de la philosophie des émotions qui ne définissent pas uniquement les états affectifs par un rôle fonctionnel mais aussi par leur constitution interne. Ces théories analysent les états affectifs, notamment les émotions, en composantes qu'elles articulent les unes aux autres : cela leur permet de se prononcer substantiellement sur la nature des états affectifs et de leurs relations, et de leur rôle dans la vie psychologique. Ces théories peuvent donc fournir au modèle cognitif de la dépression un cadre de compréhension de ce qu'il peut y avoir de dysfonctionnel dans l'humeur dépressive et ses relations avec les émotions. Je vais dégager trois cadres heuristiques : un dans lequel la relation de l'humeur et des émotions est causale, un dans lequel cette relation est logique, et un dans lequel elle est constitutive.

2. 3. Trois cadres heuristiques issus de la philosophie des émotions pour la relation dysfonctionnelle entre humeur dépressive et émotions

Les théories analytiques mettent en évidence cinq composantes principales des émotions et des humeurs (e.g. Deonna & Teroni, 2012; Scarantino & de Sousa, 2018) : 1) une composante phénoménologique (expérience, vécu, ressenti (subjectif), « éprouvé subjectif » en clinique), 2) une composante somatique (activation neurophysiologique, réponse centrale et périphérique, réponse autonome), 3) une composante intentionnelle (contenu de représentation ou attitude ou mode intentionnel d'ordre évaluatif), 4) une composante cognitive (ou mentale, états ou processus cognitifs ou mentaux), 5) une composante motivationnelle (force de motivation envers des comportements, volontaires ou involontaires, dont certains sont expressifs, et d'autres sont des actions expliquées ou justifiées par l'état affectif, et éventuellement d'autres états intentionnels). Les théories se différencient selon la composante qui est jugée fondamentale et ses articulations avec les autres composantes. En développant pour chacune l'une des principales théories existantes, je vais dégager trois cadres heuristiques pour le modèle cognitif de la dépression, en particulier pour les relations qu'il décrit entre humeur dépressive et émotion. Les théories somatiques, cognitives et motivationnelles aboutissent à définir l'humeur et les émotions par des états du corps : leurs relations sont alors des relations causales entre ces états corporels. Les théories intentionnalistes définissent les émotions et les humeurs par des contenus de représentation : leurs relations sont des relations logiques entre ces contenus. Enfin les théories phénoménologiques définissent les émotions et les humeurs par leurs qualités phénoménales : leurs relations sont des relations constitutives entre ces qualités phénoménales.

- L'humeur et les émotions comme états du corps : une relation causale

La principale théorie définissant les émotions et les humeurs par leur composante somatique, héritière de la théorie James-Lange, est celle de Prinz. Comme on l'a vu, c'est un type d'état du corps (au sens large incluant les états neuraux) qui est distinctif des émotions ou respectivement des humeurs ; un état du corps particulier de ce type est distinctif de chaque émotion particulière ou humeur particulière ; ces états corporels représentent des états des relations de l'organisme et de l'environnement. Puisque les émotions ont pour objet des choses ou des situations particulières, et les humeurs des choses ou des situations générales, elles se différencient par le fait de représenter respectivement un schème relationnel « local » ou « global » : il s'agit là d'un contenu de représentation qui exprime le fait qu'elles répondent à des objets et des événements isolés pour les émotions, et à la situation générale de l'individu dans le monde pour les humeurs. Prinz dit que les émotions et les humeurs se distinguent par leur objet ontique, défini comme la catégorie ontologique de leur type d'objets (Prinz, 2004, p. 185)

Dans la mesure où ce contenu de représentation est véhiculé par l'état du corps, les émotions sont des perceptions d'états du corps provoqués par des situations particulières. Il est plus difficile de comprendre quels sont les déterminants des états corporels propres aux humeurs, qui représentent la position générale de l'individu dans le monde : pourquoi l'état corporel de l'individu

représenterait-il sa position ou situation générale plutôt que la situation particulière dans laquelle il se trouve ? On peut envisager des causes exogènes de cet état corporel, résidant dans des situations particulières qui ont dès lors la propriété de faire signe vers la situation générale de l'individu, par exemple parce qu'elles l'illustrent ou la modifient. On peut aussi envisager des déterminants endogènes de cet état corporel, résidant dans des processus de l'organisme qui font que c'est plutôt un tel état qui est causé qu'un état typique d'une émotion dans une situation particulière. Si une émotion et une humeur se combinent ou se succèdent, par exemple si j'ai une émotion au cours d'une humeur, ou si je suis d'une certaine humeur suite à une émotion, l'état corporel dans lequel j'étais initialement va faire partie des déterminants endogènes de mon nouvel état corporel, celui dont la perception constitue mon nouvel état affectif. Les rapports entre émotions et humeurs sont donc en définitive des rapports de causalité entre états corporels.

Une telle théorie somatique nous donne alors un cadre heuristique possible pour comprendre le vécu dépressif. Supposons que l'on mette en évidence un état corporel propre à l'humeur dépressive et qu'il défavorise la formation d'états corporels typiques des émotions. Cela rendrait compte du fait que l'humeur dépressive défavorise les émotions, et que sa persistance chez le sujet dépressif prévienne la formation d'émotions chez ce dernier dans un grand nombre de situations par rapport à un sujet non dépressif exposé aux mêmes situations. Alternativement si certaines émotions négatives consistent dans des états corporels susceptibles de causer l'état corporel propre à l'humeur dépressive alors ces émotions favorisent la formation d'une humeur dépressive. Comment caractériser un dysfonctionnement de la vie affective dans ce cadre ? Puisque les émotions doivent représenter, c'est-à-dire indiquer de manière fiable, un certain état des relations de l'individu avec l'environnement, et les humeurs sa situation générale dans l'existence, ce dysfonctionnement peut être conçu en termes de désajustement entre ces relations ou cette situation générale effectives et les états corporels dont la perception constitue les émotions et humeurs de l'individu. Par exemple si l'état corporel propre à l'humeur dépressive se maintient, sa perception indique à l'individu sa position générale dans l'existence sous un jour défavorable ou indifférent, dont il ne peut pas se dégager tant que cet état est présent. De plus cet état corporel prévient la formation de certaines émotions qui se seraient formées dans certaines situations, représentant alors un certain état des relations du sujet avec son environnement, ce qui l'aurait informé sur ces relations, alors qu'en l'absence de ces émotions il ne va pas disposer de ces informations. Par exemple, si je suis d'humeur dépressive et que cela se traduit par un état de mon corps qui l'empêche d'évoluer vers un état typique de la joie dans une situation où d'autres facteurs auraient normalement provoqué cet état corporel, cela peut m'empêcher de ressentir de la joie alors qu'on vient m'annoncer le résultat d'une élection à l'égard duquel ma joie m'aurait permis de savoir que ce résultat me permettait d'anticiper une amélioration de mes droits sociaux donc de ma situation personnelle. Une théorie somatique permet ainsi de rendre compte d'un dysfonctionnement de la vie affective dans le vécu dépressif comme sous-occurrence d'émotions inhibées causalement, en tant qu'états corporels, par la présence d'un état corporel propre à l'humeur dépressive.

Une théorie cognitive et une théorie motivationnelle conduisent au même cadre heuristique.

Dans une théorie cognitive comme celle de Deonna et Teroni, une émotion est définie comme l'évaluation d'une situation et une humeur n'a pas de contenu cognitif déterminé (J. Deonna & Teroni, 2012). Elles sont l'une et l'autre des états affectifs, parce qu'elles ont une phénoménologie distinctive, une valence et une force de motivation, or cela provient du fait qu'elles font intervenir un état du corps. Leurs qualités phénoménales sont celles de cet état corporel, qui est un état de préparation à l'action (J. A. Deonna & Teroni, 2009a). Ce qui distingue l'émotion d'un autre type d'état cognitif est cet état du corps, auquel son contenu évaluatif est lié, et c'est également l'état du corps qui va définir l'humeur. L'émotion peut être définie comme une « attitude évaluative incarnée », qui est le ressenti d'un état du corps comme préparé à certaines actions, un ressenti qui est une conscience « naïve » et non réflexive de cet état du corps, qui ne le vise pas

comme tel, mais vise à travers lui ce envers quoi il nous prépare à agir (J. A. Deonna & Teroni, 2009a). Mais puisque l'humeur n'a pas de cible, qu'elle ne vise pas une situation particulière ou un objet particulier envers lequel elle préparerait à certaines actions, elle ne peut pas être définie comme une évaluation : comme elle n'est pas définie par un contenu cognitif, elle n'est pas intentionnelle (J. A. Deonna & Teroni, 2009b) donc elle consiste en définitive dans l'état du corps dont elle tire sa phénoménologie. Selon ces auteurs les rapports entre les humeurs et les émotions sont uniquement causaux : les humeurs permettent d'expliquer causalement des comportements ou des états mentaux, en particulier des émotions, mais pas de leur donner des raisons (J. A. Deonna & Teroni, 2009b). Dans une telle théorie cognitive, les humeurs consistent donc dans des états ou processus corporels qui peuvent causer ou déterminer des émotions qui sont des attitudes corporelles : les unes et les autres sont causalement reliées, et on retrouve le même cadre heuristique pour le dysfonctionnement du vécu dépressif que dans une théorie somatique.

Dans une théorie motivationnelle comme celle de Frijda (1986) les émotions sont définies comme des programmes de préparation à l'action, qui sont une activation du corps en vue de certaines tendances à l'action déterminées, s'accompagnant du ressenti subjectif propre aux états affectifs et éventuellement d'une évaluation de la situation à laquelle elles réagissent ; les humeurs sont des modes de préparation à l'action, dans lesquels il y a une activation du corps mais générale, pas en vue de tendances à l'action déterminées, les humeurs ne réagissant à aucune situation particulière. On retrouve donc un cadre dans lequel les humeurs font partie des déterminants causaux des émotions, leurs relations causales reliant des états corporels (des états d'activation ou de mobilisation du corps) qui sous-tendent les unes et les autres.

- L'humeur et les émotions comme états représentationnels : une relation logique

Une théorie intentionnaliste des émotions et des humeurs, qui les définit par un contenu de représentation, donne un cadre heuristique différent pour les relations entre émotions et humeurs, car elles entretiennent alors des rapports logiques. Les composantes qu'elles ont en tant qu'états affectifs (l'état du corps, les états cognitifs, la force de motivation, les qualités phénoménales) sont contingentes à l'égard de leur contenu de représentation, considéré comme ce qu'il y a d'essentiel dans chacune d'elles. Les émotions se distinguent alors de l'humeur par le type de ce contenu de représentation qu'elles manifestent. Une telle théorie intentionnaliste peut être naturaliste ou non naturaliste selon que le contenu de représentation des états affectifs survient ou non sur des propriétés naturelles (physiques, matérielles).

Une théorie intentionnaliste « naturaliste » importante est défendue par Michael Tye, dans le cadre d'une théorie représentationnaliste de la conscience (Tye, 1995). Puisqu'il faut rendre compte du fait qu'une émotion vise à la fois un objet ou un état du monde, et un état du corps, il faut que les deux soient mis en relation au sein de son contenu de représentation : un objet est donc représenté comme ayant certaines qualités affectives ou évaluatives en vertu desquelles il cause un certain état du corps (Tye, 2008). Dans l'humeur l'objet visé est le monde entier. L'expérience émotionnelle met donc en relation une expérience perceptuelle dirigée vers un objet (quelque chose d'extérieur) et l'expérience d'un certain état du corps, l'objet étant visé comme la cause de l'état, un contenu conceptuel pouvant s'adjoindre au contenu non-conceptuel dans lequel consistent la perception de l'objet et l'état du corps. Le contenu phénoménal de l'expérience émotionnelle articule le ressenti associé à l'expérience perceptuelle et celui associé à l'expérience corporelle par un ressenti de la connexion causale qui existe entre les deux. Son contenu représentationnel consiste dans cette double expérience perceptuelle et corporelle puisque les propriétés de l'état perceptuel sont corrélées aux qualités évaluatives de l'objet perçu, et les propriétés de l'état corporel sont visées comme causées par celui-ci. Pour concevoir comment les qualités évaluatives ou affectives sont liées aux qualités physiques ou matérielles de l'objet de l'émotion, par lesquelles celui-ci cause l'état du corps, Tye dit qu'elles surviennent dessus (Tye, 2008).

La relation entre une émotion et une humeur est alors une relation logique dans laquelle les qualités évaluatives qu'elles représentent se combinent pour former un contenu de représentation qui est la résultante de leurs contenus de représentations respectifs. Si on est d'une certaine humeur c'est qu'on se représente le monde comme ayant certaines qualités évaluatives en même temps qu'un état de son corps causé par les qualités matérielles du monde sur lesquelles ces qualités évaluatives surviennent. Si on rencontre alors un objet qui suscite en nous une émotion, on se représente les qualités évaluatives de cet objet comme survenant sur des qualités matérielles par lesquelles celui-ci cause l'état de notre corps. Comme cet état était déjà en partie déterminé par des qualités du monde sur lesquelles survenaient les qualités évaluatives que représentait notre humeur, on peut dire que les qualités évaluatives que notre émotion nous représente liées à l'objet qui cause l'état du corps se combinent avec celles que l'humeur nous représente liées au monde entier. Autrement dit, si une émotion survient lorsque nous étions déjà affectés par une humeur, les qualités évaluatives ou affectives que nous nous représentons dans l'objet de cette émotion en lien avec les qualités corporelles de l'état du corps que cause cet objet, vont être différentes de ce qu'elles auraient été en l'absence de cette humeur puisque nous nous serions alors trouvés dans un état corporel différent. Comme l'émotion est définie par son contenu de représentation, donc par les qualités évaluatives de cet objet, et que celles-ci sont en partie déterminées par les qualités évaluatives du monde entier que représente l'humeur déjà présente avant l'émotion, leur relation est celle d'une combinaison de propriétés évaluatives, soit une relation logique.

Dans la théorie intentionnaliste non « naturaliste » de Christine Tappolet, qui défend un réalisme perceptuel, les émotions sont définies comme des perceptions de valeurs. Les valeurs sont des propriétés axiologiques monadiques instanciées par leurs objets, que l'expérience émotionnelle nous fait connaître grâce à ses qualités affectives. Nous décrivons ces qualités par les concepts évaluatifs et affectifs par lesquels nous qualifions les objets de nos émotions, par exemple comme « effrayants » ou « dégoûtants ». (Tappolet, 2000, 2016) Les humeurs perçoivent des propriétés évaluatives potentielles, et non actuelles, donc présentées par l'état affectif sans être instanciées par un objet (Tappolet, 2000, 2018). Plus précisément les émotions nous permettent de connaître les propriétés axiologiques de leurs objets à condition d'être des émotions correctes, ce qui nécessite qu'elles se déroulent dans des conditions normales, où n'intervient aucune circonstance susceptible de perturber ce déroulement, d'interférer avec notre expérience émotionnelle, circonstances que Tappolet appelle « *defeater* » (Tappolet, 2016). L'humeur est alors un « *defeater* » de l'émotion, c'est-à-dire qu'elle fait partie des facteurs qui peuvent biaiser ou fausser une émotion : cela signifie que les propriétés évaluatives que représente l'humeur vont déterminer ou modifier celles que nous percevons dans l'objet de l'émotion, voire nous faire percevoir un contenu axiologique erroné, des qualités évaluatives monadiques que cet objet n'instancie pas. La relation entre humeur et émotions est également une combinaison de contenus de représentation qui sont des propriétés évaluatives, donc c'est une relation logique.

Ces théories intentionnalistes donnent donc un cadre heuristique au modèle cognitif de la dépression en termes de relations logiques entre contenus de représentation. Plus précisément comment ces théories permettent-elles de comprendre les relations décrites par ce modèle entre humeur dépressive, émotions et cognitions ? Les émotions sont définies par des contenus qui sont la représentation de qualités évaluatives d'objets particuliers, alors que les humeurs représentent des qualités évaluatives du monde. La présence d'une humeur peut déterminer le fait que l'on ait plus ou moins certaines émotions en déterminant les qualités évaluatives que l'on est susceptible de se représenter comme instanciées par des objets particuliers compte tenu des qualités évaluatives qu'elle représente comme inhérentes au monde. Par exemple si l'humeur dépressive représente le monde comme indifférent, sa présence peut empêcher de ressentir de la colère dans une situation donnée, parce que les qualités que cette humeur représente se retrouvent dans l'objet qu'on

rencontre et que dès lors on ne se représente pas ses qualités irritantes. Les états affectifs étant une voie d'accès épistémique aux propriétés qui sont leur contenu de représentation spécifique, la présence d'une humeur dépressive représentant des qualités évaluatives négatives se traduit par la prédominance de cognitions congruentes à cette humeur, chez le sujet dépressif, car celles-ci ne sont rien d'autre que des états cognitifs ayant des contenus auxquels cette humeur donne accès. Ces états cognitifs (croyances, ruminations...) peuvent également être favorisés par des émotions ayant le même type de contenu. Enfin des émotions négatives mal régulées, si elle se traduisent par la représentation persistante de qualités évaluatives négatives, peuvent donner lieu à une humeur dépressive, définie par un tel contenu de représentation lorsque ces qualités cessent d'être instanciées par un objet puisqu'elles sont alors attribuées au monde en général.

- L'humeur et les émotions comme états phénoménaux : une relation constitutive

Si on définit les émotions et les humeurs en termes de qualités phénoménales, elles ont des relations qu'on peut qualifier de constitutives. Les qualités phénoménales par lesquelles se définit l'humeur ne s'accompagnent pas d'une visée intentionnelle, puisque l'humeur n'a pas de cible, alors que les qualités phénoménales de l'émotion accompagnent la visée intentionnelle de l'objet particulier de cette émotion. Puisque les qualités phénoménales sont la composante fondamentale, essentielle, des états affectifs, la visée intentionnelle de l'émotion se constitue à partir d'elles. En présence d'une humeur, les émotions vont donc être en partie constituées par l'humeur en question, leurs qualités phénoménales respectives se mêlant dans une seule et même phénoménologie qui détermine une émotion différente, avec une visée intentionnelle différente, de celle que l'on aurait eue en l'absence de l'humeur.

Ainsi dans la théorie de Goldie, les émotions sont définies par leur phénoménologie, par leurs qualités phénoménales qui constituent un « *feeling* » mêlant un « *feeling toward* » un objet particulier et un « *bodily feeling* » qui est le ressenti d'un état corporel (Goldie, 2000, Chapitre 3). La différence entre une émotion et une humeur est alors une différence de degré, car le « *feeling toward* » de l'humeur vise un objet plus général que celui de l'émotion (Goldie, 2000, Chapitre 6). Une émotion peut donner lieu à une humeur si ses qualités phénoménales se maintiennent mais que son objet particulier cesse d'être visé ; tout autant une humeur peut donner lieu à une émotion si la partir de l'état phénoménal dans lequel elle consiste sur forme une visée intentionnelle d'un objet particulier (Goldie, 2000, p. 19). Dans une telle théorie, les qualités phénoménales par lesquelles se distingue une humeur peuvent donc devenir constitutives d'une émotion et réciproquement.

Cela nous donne un cadre heuristique pour le vécu dépressif dans lequel l'humeur dépressive peut se constituer à partir d'émotions négatives, ou bien cette humeur dépressive peut rendre compte d'une baisse des émotions parce que sa présence exclut de l'expérience subjective les qualités phénoménales à partir desquelles pourraient se constituer les visées intentionnelles propres à ces émotions. Par exemple une émotion prolongée de tristesse, consistant dans une certaine visée intentionnelle d'un objet, inhérente à des qualités phénoménales proches de celles de l'humeur dépressive, peut donner lieu à une humeur dépressive si cette visée devient générale ou absente alors que les qualités phénoménales persistent. Inversement la présence d'une humeur dépressive consistant dans des qualités phénoménales proches de celles de la tristesse peut donner lieu à une telle émotion dans une situation où la visée intentionnelle d'un certain objet va se constituer à partir de ces qualités phénoménales. Ou bien la persistance d'une humeur dépressive consistant dans des qualités phénoménales différentes ferait qu'aucune visée intentionnelle caractéristique d'une émotion donnée ne pourrait se constituer à partir de ces qualités dans des situations qui auraient normalement suscité ces émotions : ce serait un bon cadre heuristique pour comprendre le modèle de l'insensibilité émotionnelle au contexte comme définissant la dépression par un dysfonctionnement de la vie psychologique car cela veut dire que l'expérience subjective est dominée par des qualités phénoménales à partir desquelles ne se constituent pas certaines visées intentionnelles, qui se seraient constituées en l'absence de cette humeur.

- Albert, P. R., & Lemonde, S. (2004). 5-HT1A receptors, gene repression, and depression : Guilt by association. *The Neuroscientist*, *10*(6), 575-593.
- Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology : A meta-analytic review. *Clinical psychology review*, *30*(2), 217-237.
- American psychiatric association. (2013). *DSM-5 : Diagnostic and statistical manual of mental disorders / American Psychiatric Association* (5th edition). American Psychiatric Publishing.
- Andrews, P. W., & Thomson Jr, J. A. (2009). The bright side of being blue : Depression as an adaptation for analyzing complex problems. *Psychological review*, *116*(3), 620.
- Azorin, J. M., & Kaladjian, A. (2009). Dépression et rythmes circadiens. *L'Encéphale*, *35*, S68-S71.
- Baeyens, C. (2019). Les facteurs cognitifs dans la régulation émotionnelle. *Neuropsychologie en psychiatrie*, 157-176.
- Barkow, J. H., Cosmides, L., & Tooby, J. (1992). *The adapted mind : Evolutionary psychology and the generation of culture*. Oxford University Press, USA.
- Barrett, L. F. (2006). Are Emotions Natural Kinds? *Perspectives on Psychological Science*, *1*(1), 28-58. <https://doi.org/10.1111/j.1745-6916.2006.00003.x>
- Beck, A. T. (1967). *Depression. Clinical, Experimental and Theoretical Aspects*. New York (Hoeber) 1967.
- Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory : Twenty-five years of evaluation. *Clinical Psychology Review*, *8*(1), 77-100. [https://doi.org/10.1016/0272-7358\(88\)90050-5](https://doi.org/10.1016/0272-7358(88)90050-5)
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*, 561-571. <https://doi.org/10.1001/archpsyc.1961.01710120031004>
- Bergner, R. M. (1997). What is psychopathology? And so what? *Clinical Psychology: Science and Practice*, *4*(3), 235-248.
- Bolton, D. (2008). *What is Mental Disorder? : An Essay in Philosophy, Science, and Values*. Oxford University Press.

- Bourke, C., Douglas, K., & Porter, R. (2010). Processing of facial emotion expression in major depression : A review. *Australian & New Zealand Journal of Psychiatry*, 44(8), 681-696.
- Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., & Mancill, R. B. (2001). Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *Journal of abnormal psychology*, 110(4), 585.
- Carleton, R. N., Thibodeau, M. A., Teale, M. J., Welch, P. G., Abrams, M. P., Robinson, T., & Asmundson, G. J. (2013). The center for epidemiologic studies depression scale : A review with a theoretical and empirical examination of item content and factor structure. *PloS one*, 8(3), e58067.
- Castel, P.-H. (2009). *L'esprit malade : Cerveaux, folies, individus*. Ithaque.
- Chalmers, D. J. (2002). *Philosophy of mind : Classical and contemporary readings*.
- Cléry-Melin, M.-L., Schmidt, L., Lafargue, G., Baup, N., Fossati, P., & Pessiglione, M. (2011). Why don't you try harder? An investigation of effort production in major depression. *PloS one*, 6(8), e23178.
- Colombel, F. (2007). Biais mnésique associé à l'humeur dépressive : Quelques réflexions critiques. *L'encéphale*, 33(3), 242-248.
- Colombetti, G. (2009). From affect programs to dynamical discrete emotions. *Philosophical Psychology*, 22(4), 407-425.
- Corruble, E. (2010). Dépression : Concepts actuels. *L'Encéphale*, 36, S104-S107.
- Cummins, R. (1975). Functional Analysis. *Journal of Philosophy*, 72(November), 741-764.
<https://doi.org/10.2307/2024640>
- Darwin, C. (1877). *L'expression des émotions chez l'homme et les animaux*.
- Demazeux, S. (2013). *Qu'est-ce que le DSM? Genèse et transformations de la bible américaine de la psychiatrie*. Ithaque.
- Deonna, J. A., & Teroni, F. (2009a). L'intentionnalité des émotions : Du corps aux valeurs. *Revue européenne des sciences sociales. European Journal of Social Sciences*, XLVII-144, 25-41.
- Deonna, J. A., & Teroni, F. (2009b). Taking affective explanations to heart. *Social Science Information*, 48(3), 359-377.
- Deonna, J., & Teroni, F. (2012). *The emotions : A philosophical introduction*. Routledge.

- Desseilles, M. (2010). Les modèles étiologiques de la dépression. *Acta psychiatrica belgica*, 110(4), 34-42.
- Ehring, T., Tuschen-Caffier, B., Schnülle, J., Fischer, S., & Gross, J. J. (2010). Emotion regulation and vulnerability to depression : Spontaneous versus instructed use of emotion suppression and reappraisal. *Emotion*, 10(4), 563.
- Faucher, L. (2016). Darwinian blues : Evolutionary psychiatry and depression. In *Sadness or Depression?* (p. 69-94). Springer.
- Faucher, L. (2021). 3 Facts, Facts, Facts : HD Analysis Goes Factual. *Defining Mental Disorder: Jerome Wakefield and His Critics*, 47.
- Faucher, L., & Forest, D. (2021). *Defining mental disorder : Jerome Wakefield and his critics*. MIT Press.
- Faucher, L., & Goyer, S. (2015). RDoC : Thinking outside the DSM box without falling into a reductionist trap. In *The DSM-5 in perspective* (p. 199-224). Springer.
- Forest, D. (2016). Is an Anatomy of Melancholia Possible? Brain Processes, Depression, and Mood Regulation. In *Sadness or Depression?* (p. 95-107). Springer.
- Forest, D., & Le Bidan, M. (2016). In Search of Normal Functions : BST, Cummins Functions, and Hempel's Problem. In *Naturalism in the Philosophy of Health* (p. 39-51). Springer.
- Frijda, N. H. (1986). *The emotions*. Cambridge University Press.
- Gaillard, R., Gourion, D., & Llorca, P. M. (2013). L'anhédonie dans la dépression. *L'encephale*, 39(4), 296-305.
- Goldie, P. (2000). *The emotions : A philosophical exploration*. Oxford University Press.
- Griffiths, P. E. (1996). The historical turn in the study of adaptation. *The British journal for the philosophy of science*, 47(4), 511-532.
- Griffiths, P. E. (1997). *What emotions really are : The problem of psychological categories*. University of Chicago Press.
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes : Implications for affect, relationships, and well-being. *Journal of personality and social psychology*, 85(2), 348.

- Guilbaud, O., Corcos, M., Hjalmarsson, L., Loas, G., & Jeammet, P. (2002). Modèles psychoneuro-immunologiques de la dépression : Une revue critique. *Annales Médico-psychologiques, revue psychiatrique*, *160*(2), 179-185.
- Hamilton, M. (1986). The Hamilton rating scale for depression. In *Assessment of depression* (p. 143-152). Springer.
- Hautgen, T. (2014). Jules Séglas ou la première formulation moderne des états dépressifs. *PSN*, *12*(4), 47-65.
- Horwitz, A. V., & Wakefield, J. C. (2007). *The loss of sadness : How psychiatry transformed normal sorrow into depressive disorder*. Oxford University Press.
- Joormann, J. (2010). Cognitive inhibition and emotion regulation in depression. *Current Directions in Psychological Science*, *19*(3), 161-166.
- Joormann, J., & Gotlib, I. H. (2006). Is this happiness I see? Biases in the identification of emotional facial expressions in depression and social phobia. *Journal of abnormal psychology*, *115*(4), 705.
- Joormann, J., & Gotlib, I. H. (2010). Emotion regulation in depression : Relation to cognitive inhibition. *Cognition and Emotion*, *24*(2), 281-298.
- Kendler, K. S., Gardner, C. O., & Prescott, C. A. (2002). Toward a comprehensive developmental model for major depression in women. *American Journal of Psychiatry*, *159*(7), 1133-1145.
- Kendler, K. S., Gardner, C. O., & Prescott, C. A. (2006). Toward a comprehensive developmental model for major depression in men. *American Journal of Psychiatry*, *163*(1), 115-124.
- Leppänen, J. M. (2006). Emotional information processing in mood disorders : A review of behavioral and neuroimaging findings. *Current opinion in psychiatry*, *19*(1), 34-39.
- Lesch, K. P. (2004). Gene–environment interaction and the genetics of depression. *Journal of Psychiatry and Neuroscience*, *29*(3), 174.
- Martin, R. C., & Dahlen, E. R. (2005). Cognitive emotion regulation in the prediction of depression, anxiety, stress, and anger. *Personality and individual differences*, *39*(7), 1249-1260.
- Mayberg, H. S. (2007). Defining the Neural Circuitry of Depression : Toward a New Nosology With Therapeutic Implications. *Biological Psychiatry*, *61*(6), 729-730.
<https://doi.org/10.1016/j.biopsych.2007.01.013>

- Miller, I. W., Bishop, S., Norman, W. H., & Maddever, H. (1985). The modified Hamilton rating scale for depression : Reliability and validity. *Psychiatry Research*, *14*(2), 131-142.
[https://doi.org/10.1016/0165-1781\(85\)90057-5](https://doi.org/10.1016/0165-1781(85)90057-5)
- Millikan, R. G. (1989). In Defense of Proper Functions. *Philosophy of Science*, *56*(2), 288-302.
<https://doi.org/10.1086/289488>
- Mirabel-Sarron, C. (2012). Dysrégulations émotionnelles et dépressions. *L'Encéphale*, *38*, S41-S44.
- Mouchabac, S. (2006). La dépression : Évolution ou révolution? Une approche évolutionniste des troubles de l'humeur. *Neuropsychiatrie: tendances et débats*, *29*, 19-36.
- Neander, K. (1991). Functions as Selected Effects : The Conceptual Analyst's Defense. *Philosophy of Science*, *58*(2), 168-184. <https://doi.org/10.1086/289610>
- Nesse, R. M. (2009). Explaining depression : Neuroscience is not enough, evolution is essential. *Understanding depression: A translational approach*, 17-35.
- Nesse, R. M., & Williams, G. C. (1998). Evolution and the origins of disease. *Scientific American*, *279*(5), 86-93.
- Nesse, R. M., Williams, G. C., & Mysterud, I. (1995). Why we get sick. *Trends in Ecology and Evolution*, *10*(7), 300-301.
- Nettle, D. (2004). Evolutionary origins of depression : A review and reformulation. *Journal of affective disorders*, *81*(2), 91-102.
- Panksepp, J. (2003). Feeling the pain of social loss. *Science*, *302*(5643), 237-239.
- Pasquier, A. (2012). *Psychopathologie des émotions*. Armand Colin.
- Peeters, F., Berkhof, J., Rottenberg, J., & Nicolson, N. A. (2010). Ambulatory emotional reactivity to negative daily life events predicts remission from major depressive disorder. *Behaviour research and therapy*, *48*(8), 754-760.
- Polosan, M., Lemogne, C., Jardri, R., & Fossati, P. (2016). La cognition au cœur de la dépression. *L'Encéphale*, *42*(1), 1S3-1S11.
- Price, J., Sloman, L., Gardner, R., Gilbert, P., & Rohde, P. (1994). The social competition hypothesis of depression. *The British Journal of Psychiatry*, *164*(3), 309-315.
- Prinz, J. J. (2004). *Gut reactions : A perceptual theory of emotion*. Oxford University Press.

- Radden, J. (2009). *Moody minds distempered : Essays on melancholy and depression*. Oxford University Press on Demand.
- Rottenberg, J. (2005). Mood and emotion in major depression. *Current Directions in Psychological Science*, 14(3), 167-170.
- Rottenberg, J., Gross, J. J., & Gotlib, I. H. (2005). Emotion context insensitivity in major depressive disorder. *Journal of abnormal psychology*, 114(4), 627.
- Scarantino, A., & de Sousa, R. (2018). *Emotion*.
- Sharp, R. (2015). The Hamilton Rating Scale for Depression. *Occupational Medicine*, 65(4), 340-340. <https://doi.org/10.1093/occmed/kqv043>
- Siemer, M. (2009). Mood Experience : Implications of a Dispositional Theory of Moods. *Emotion Review*, 1(3), 256-263. <https://doi.org/10.1177/1754073909103594>
- Sizer, L. (2000). Towards a computational theory of mood. *The British Journal for the Philosophy of Science*, 51(4), 743-770.
- Surguladze, S. A., Young, A. W., Senior, C., Brébion, G., Travis, M. J., & Phillips, M. L. (2004). Recognition accuracy and response bias to happy and sad facial expressions in patients with major depression. *Neuropsychology*, 18(2), 212.
- Tappolet, C. (2000). *Emotions et valeurs*. FeniXX.
- Tappolet, C. (2016). *Emotions, values, and agency*. Oxford University Press.
- Tappolet, C. (2018). The metaphysics of moods. In *The ontology of emotions* (p. 169-186). Cambridge University Press.
- Tye, M. (1995). *Ten problems of consciousness : A representational theory of the phenomenal mind*. MIT Press.
- Tye, M. (2008). The Experience of Emotion : An Intentionalist Theory. *Revue Internationale de Philosophie*, n° 243(1), 25-50.
- Wakefield, J. C. (1992). The concept of mental disorder : On the boundary between biological facts and social values. *American psychologist*, 47(3), 373.
- Widlöcher, D. (2002). Dépression et anxiété. *Revue française de psychanalyse*, 66(2), 409-422.