

La nature des émotions : une recension partisane

Constant Bonard

Volume 18, numéro 1, 2024

URI : <https://id.erudit.org/iderudit/1117228ar>

DOI : <https://doi.org/10.7202/1117228ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Centre de recherche en éthique (CRÉ)

ISSN

1718-9977 (numérique)

[Découvrir la revue](#)

Citer ce document

Bonard, C. (2024). La nature des émotions : une recension partisane. *Les ateliers de l'éthique / The Ethics Forum*, 18(1), 29–49.
<https://doi.org/10.7202/1117228ar>

Résumé de l'article

Dans cette recension, je proposerai un résumé ainsi que des commentaires critiques, chapitre par chapitre, du livre de Samuel Lepine *La nature des émotions : Une introduction partisane* (2023). Je ferai un rapide compte-rendu des premiers chapitres, mais discuterai plus en détail les trois derniers, où se situent les contributions les plus inédites et, à mes yeux, les plus intéressantes.



LA NATURE DES ÉMOTIONS : UNE RECENSION PARTISANE

CONSTANT BONARD
UNIVERSITÉ DE BERNE

RÉSUMÉ :

Dans cette recension, je proposerai un résumé ainsi que des commentaires critiques, chapitre par chapitre, du livre de Samuel Lepine *La nature des émotions : Une introduction partisane* (2023). Je ferai un rapide compte-rendu des premiers chapitres, mais discuterai plus en détail les trois derniers, où se situent les contributions les plus inédites et, à mes yeux, les plus intéressantes.

ABSTRACT:

In this review, I will offer a summary as well as critical comments, chapter by chapter, of Samuel Lepine's book *La nature des émotions: Une introduction partisane* [The nature of emotions: An opinionated introduction] (2023). I will briefly summarize the first few chapters and discuss in more detail the last three chapters, where the most novel and most interesting contributions of the author are to be found.

INTRODUCTION

J'offre ici une recension de l'excellent livre de Samuel Lepine *La nature des émotions : Une introduction partisane* (2023). Il s'agit d'une introduction aux grandes questions de la philosophie analytique sur les émotions. Le livre présente également de nombreux travaux pertinents en psychologie et en philosophie des sciences. Mais cet ouvrage va au-delà de sa mission introductive : il contient plusieurs contributions originales de l'auteur. L'ouvrage se concentre sur cette hypothèse : l'idée que « les émotions jouent un rôle crucial dans notre capacité à prendre en compte la valeur que les choses ou les personnes peuvent avoir pour nous » (Lepine, 2023, p. 9). Bien que cette hypothèse soit partagée par plusieurs philosophes des émotions, elle est détaillée de façon originale par l'auteur, prenant au fur et à mesure du livre la forme de la thèse dite « motivationnelle » propre à Samuel Lepine, laquelle sera défendue plus spécifiquement dans les deux derniers chapitres.

Je recommande vivement ce livre, à la fois abordable et pointu, autant aux néophytes dans ce domaine qu'aux spécialistes. L'ouvrage de Samuel Lepine n'a rien à envier aux livres publiés sur ce sujet par les meilleures presses internationales.

CHAPITRE I : UNE ANALYSE NAÏVE

Ce premier chapitre brosse les questions principales de la philosophie des émotions sans forcément fournir de réponses, sans encore s'engager dans les diverses théories proposées par les philosophes contemporains – d'où son sous-titre. L'auteur commence par présenter certaines caractéristiques typiquement attribuées aux émotions tout en introduisant les termes techniques utilisés au long de l'ouvrage (par exemple : automaticité, valence, intentionnalité, base cognitive, base motivationnelle). Le chapitre I distingue également les émotions d'autres types d'états affectifs : les humeurs, sentiments, passions, affects, et désirs. Enfin, il présente et discute la question « Y a-t-il des émotions inconscientes ? », à laquelle il répond par la négative, sur la base de certaines études scientifiques et de la taxonomie présentée auparavant.

Pour émettre un regard critique sur ce chapitre initial, je noterai que, bien qu'elle soit qualifiée de « naïve », l'analyse proposée inscrit l'ouvrage dans un sillon assez spécifique de la philosophie des émotions en faisant certains présupposés théoriques qui ne sont pas toujours consensuels. Je pense aux hypothèses suivantes, que je me contenterai d'énoncer : que la conscience phénoménale est essentielle à un épisode émotionnel (contraire à Prinz, 2005 ; Scarantino, 2014) ; les émotions ont (uniquement) une direction d'ajustement esprit-monde (contraire à Whiting, 2011 ; Scarantino, 2014) ; que ce sont les émotions, et non les états mentaux qui précèdent ou fondent celles-ci, qui permettent d'évaluer les objets sur lesquels elles portent (contraire à Mulligan, 2009 ; Müller, 2019). S'agissant de la question des émotions inconscientes, on peut noter également que certaines études scientifiques (pour des recensions, voir Winkielman et

Berridge, 2004 ; Smith et Lane, 2016) et philosophiques (Prinz, 2005 ; Scarantino, 2014) qui postulent l'existence d'émotions inconscientes ne sont pas discutées. Mais, par ailleurs, le chapitre joue parfaitement son rôle d'initiation à la philosophie contemporaine des émotions – ce qui est le plus important à ce stade de l'ouvrage.

CHAPITRE II : LES ÉMOTIONS ET LA COGNITION

Le chapitre II porte sur les relations entre émotion et cognition. Cette dernière est comprise comme l'ensemble des mécanismes mentaux dont le rôle est de traiter l'information et qui rendent ainsi les stimuli disponibles pour le traitement sémantique (Lepine, 2023, p. 61). Ce chapitre ne ressemble en rien aux introductions en philosophie des émotions que je connais et donne une appréciable impression de fraîcheur au sein de cette littérature.

La question de la relation entre émotion et cognition n'est pas seulement abordée sous l'angle de la philosophie – l'auteur discute entre autres les idées de Fodor ou Prinz sur la question –, le chapitre présente également de façon remarquable l'histoire de cette question en psychologie. Il part de la théorie des émotions de James, présente la théorie bifactorielle de Schachter et Singer, retrace le fameux débat entre Zajonc et Lazarus, et présente la théorie de l'évaluation (*appraisal theory*). Cette dernière est la théorie psychologique que privilégie l'auteur, et il explique pourquoi dans ce chapitre. Elle lui permet d'affirmer que les émotions comportent bien une partie cognitive et qu'elles sont dotées de représentations évaluatives postulées par la théorie de l'évaluation (pour une analyse détaillée de cette conclusion, voir Bonard, 2021a, chap. 9 ; 2024).

Ce chapitre constitue une excellente introduction à la psychologie cognitive des émotions, du moins en ce qui concerne la question du chapitre : la relation entre émotion et cognition. Notons néanmoins qu'il s'agit plutôt d'une introduction à l'histoire de cette question que d'une recension des derniers développements (pour une introduction plus complète, voir Moors, 2022). Ce chapitre donne également des clefs importantes pour la philosophie des sciences cognitives, notamment par sa discussion de la définition de la cognition.

CHAPITRE III : L'UNITÉ DES ÉMOTIONS

Le chapitre III s'attaque à la question de l'unité des émotions, à savoir si celles-ci forment ou non ce qu'on appelle en philosophie des sciences et en métaphysique une *espèce naturelle*. D'un point de vue chimique, l'or est une espèce naturelle, puisque sa molécule fait partie des éléments chimiques ; tout or pur a les mêmes propriétés chimiques. Par contraste, l'air n'est pas une espèce naturelle pour les chimistes : il est un composé de différents éléments chimiques disparates dont la composition et la proportion varient d'un endroit à l'autre (azote, oxygène, gaz carbonique, etc.). Les espèces naturelles sont donc des classes cohérentes regroupant des phénomènes qui partagent une « essence commune, ou à tout le moins, des propriétés communes leur conférant des

pouvoirs causaux identiques » (Lepine, 2023, p. 72) et qui méritent ainsi que les sciences les considèrent comme un objet d'étude fécond. Les émotions possèdent-elles suffisamment de propriétés communes pour être ainsi qualifiées ? Le chapitre défend l'idée que c'est effectivement le cas et s'attaque en particulier aux arguments donnés par Griffiths (2008) contre cette thèse.

Ce chapitre est donc surtout consacré à la philosophie des sciences ainsi qu'à la métaphysique des émotions. Samuel Lepine aborde cette question en s'appuyant non seulement sur la littérature philosophique (Quine, Fodor, Kripke, Boyd, Griffiths), mais aussi sur de nombreuses études scientifiques en psychologie des émotions, en neurosciences, et en psychologie évolutionnaire.

Encore une fois, il faut louer l'auteur pour cette discussion des recherches scientifiques pertinentes aux questions philosophiques abordées. La présentation est accessible, claire, ne s'embarrasse pas de détails inutiles et fait preuve d'équité envers les différents courants scientifiques discutés. L'auteur présente en particulier le débat en psychologie des émotions entre la théorie des émotions basiques (Darwin, Ekman, Tobby et Cosmides, Nesse, entre autres) et la théorie des émotions socialement construites (Russell, Barrett, Averill en psychologie, mais aussi Doi et Lutz en sciences sociales). En gros, ce débat concerne la question suivante : les émotions sont-elles innées, encodées dans nos gènes ou, au contraire, construites socialement, changeant d'une culture à une autre ?

L'auteur concède des arguments aux deux parties, mais conclut que les faits mis en avant par les constructivistes, bien que permettant de douter de l'unité des émotions, ne sont pas suffisants pour contrer l'idée que les émotions constituent une espèce naturelle. Mais cela ne l'amène pas à épouser la théorie des émotions basiques à la Ekman. Il soutient, en lien avec les autres chapitres du livre, que chaque type d'émotions peut être défini par son « objet formel » ou son « thème relationnel », c'est-à-dire par le fait que chaque type d'émotions porte sur un ensemble d'objets unifiés par l'évaluation propre à ce type. Pour illustrer : on peut avoir peur de tout un tas de choses différentes et cet ensemble disparate change d'une culture à l'autre – toutefois, dans toute culture, les objets de la peur sont évalués par celui ou celle qui ressent la peur comme des menaces. Autre exemple : on peut s'indigner de tout et de rien, les objets de l'indignation changent selon l'époque et le lieu – toutefois, on s'indigne toujours de ce qui est perçu comme injuste. De même : toute joie porte sur quelque chose qui est vécu comme favorable. Et ainsi de suite pour chaque type d'émotions. Dans les mots de Lepine : « Autrement dit, les émotions constitueraient bien une espèce naturelle dans la mesure où elles partagent la fonction de détecter pour chacune un thème relationnel qui leur est propre [comme la menace, l'injustice, le favorable]. » (Lepine, 2023, p. 87)

Notons en passant que cette thèse est en tension avec l'idée selon laquelle au moins certaines humeurs, bien qu'étant différentes des émotions sur plusieurs points, possèdent les mêmes thèmes relationnels que les émotions auxquelles elles correspondent (Bonard, 2022 ; Rossi, 2021 ; Mitchell, 2019b).

Par exemple, une humeur triste et la tristesse partageraient le thème relationnel de la perte. L'auteur n'a pas donné d'arguments contre cette idée dans sa discussion des humeurs au premier chapitre. Peut-être serait-il prêt à accepter que les humeurs en question appartiennent aux mêmes espèces naturelles que les émotions et qu'ainsi les scientifiques puissent les étudier comme un tout ? Cela ne serait sans doute pas en contradiction avec la théorie psychologique de l'évaluation discutée au chapitre précédent.

Le restant du chapitre III traite de deux autres questions pertinentes au thème de l'unité des émotions : « Les émotions sont-elles modulaires ? » et « Y a-t-il un cerveau des émotions ? ». À la première question, l'auteur répond, de façon plausible, que les émotions ne sont que pseudo-modulaires si l'on comprend la modularité comme la définit Fodor (1983). Il concède néanmoins qu'elles peuvent être considérées comme étant modulaires au sens large donné à cette notion par la théorie de modularité massive (voir par exemple Sperber, 1996). En ce qui concerne la seconde question, l'auteur répond que, étant donné les enquêtes scientifiques discutées, il est plausible que certaines parties du cerveau soient responsables de l'évaluation des stimuli, en accord avec la théorie de l'évaluation privilégiée par Lepine lors du chapitre précédent. Les chapitres II et III se renforcent ainsi mutuellement.

CHAPITRE IV : LES THÉORIES DES ÉMOTIONS

Ayant posé les bases conceptuelles et scientifiques dans les trois premiers chapitres, Samuel Lepine consacre le reste de l'ouvrage à des questions abordées dans le débat contemporain en philosophie des émotions. Il ne fait pas que présenter ces questions, il y apporte des réponses originales, qu'il a développées au cours des dernières années, notamment dans sa thèse de doctorat consacrée aux émotions et défendue en 2016.

Dans le chapitre IV, l'auteur part du constat que les émotions sont des attitudes évaluatives, dans le sens détaillé au chapitre III : chaque émotion porte sur un thème relationnel qui est une forme d'évaluation de la situation comme étant plus ou moins positive pour le sujet – l'objet des émotions est évalué comme impliquant un danger, une perte, un succès, etc. Il discute d'un ensemble de théories philosophiques ayant détaillé cette nature évaluative des émotions ; des théories analysant en quoi, par exemple, éprouver de la peur envers un animal implique d'évaluer que cet animal constitue une menace pour soi. Sont passées en revue les théories dites judicative, mixte, quasi-perceptuelle, perceptuelle, et attitudinale. À titre d'exemple, la théorie judicative fait l'hypothèse que l'on a peur de quelque chose lorsqu'on *juge* (qu'on croit activement) que cette chose est dangereuse pour soi. La théorie perceptuelle prétend au contraire que la peur s'apparente plutôt à une *perception* de danger et non à un jugement – les deux se distinguent sous plusieurs aspects, comme le montre notamment l'illusion de Müller-Lyer où on perçoit deux formes (ressemblant à <—> et >—<) comme étant de longueurs différentes même lorsqu'on sait bien (et donc qu'on juge) qu'elles sont de même longueur.

Après avoir exposé des arguments pour et contre les différentes théories citées, l'auteur conclut que la théorie attitudinale est la meilleure en l'état actuel des connaissances (pour un compte-rendu complémentaire, voir Bonard, 2021b). Il s'agit d'une théorie originellement proposée par Julien Deonna et Fabrice Teroni (2012 – un ouvrage auquel celui de Lepine doit beaucoup). Voici les trois hypothèses majeures de cette théorie :

- Premièrement, les émotions ne peuvent pas être réduites à un autre type d'état mental ; ni à un jugement de valeur (comme le veut la théorie judiciaire), ni à une perception de valeur (comme le prétend la théorie perceptuelle), ni à aucun autre état mental portant sur une valeur. En d'autres termes, les émotions sont des attitudes évaluatives *sui generis*. Chaque type d'émotions (tristesse, colère, joie, etc.) est une attitude psychologique qui mérite sa propre analyse et qui ne peut être entièrement comprise par le biais d'autres types d'états mentaux.
- Deuxièmement, chaque émotion possède sa propre signature de sensations corporelles. Ces dernières sont comprises par la théorie attitudinale essentiellement comme la conscience des modifications automatiques de nos corps se préparant à réagir à l'objet de l'émotion : la personne ressent que son corps se prépare à s'approcher, fuir, interagir, ou agresser l'objet de l'émotion, entre autres tendances à l'action.
- Troisièmement, les émotions ne portent pas sur, ni ne représentent, des valeurs – contrairement à ce que prétendent les autres théories mentionnées ci-dessus. Ce sont des attitudes évaluatives dans le sens où, en portant sur des objets divers et variés, elles présupposent en quelque sorte une évaluation de ceux-ci en vertu de la nature de l'attitude en question. Cela serait comparable au fait que, lorsque je *crois* que l'équipe de France de basket est éliminée, je présuppose d'une certaine manière que cela est vrai, sans pour autant avoir besoin de juger activement que cette phrase est vraie. Lorsque je me souviens de l'élimination de l'équipe de France, je présuppose d'une certaine manière que cela se déroulait dans le passé, sans avoir à juger de cela activement. Et lorsque je suis triste que l'équipe de France se soit fait éliminer, je présuppose d'une certaine manière que cet événement constitue une perte pour moi, sans avoir à juger ou me représenter cela activement. La tristesse constituerait ainsi une évaluation de l'événement comme étant une perte, mais la tristesse ne porte pas directement sur la perte. Elle porte sur l'événement – tout comme la croyance porte sur l'événement et non sur la vérité (qu'elle présuppose) et comme le souvenir porte sur l'événement et non sur le fait qu'il appartient au passé (qu'il présuppose).

La théorie attitudinale de Deonna et Teroni est raffinée – voire modifiée – par Lepine, qui a recours à l'analyse des attitudes psychologiques proposée par Cummins (1996) pour éclairer en quoi les attitudes – et donc les diverses émotions – se distinguent les unes des autres. Cette analyse explicite ce que j'entends par le fait que diverses attitudes « présupposent d'une certaine manière » diverses informations. L'analyse de Cummins est fonctionnaliste : chaque atti-

tude possède une fonction particulière dans le traitement de l'information sur laquelle porte l'attitude, et c'est analyser cette fonction qui permet d'analyser l'attitude. Croire que les Bleus sont éliminés, se souvenir de l'élimination des Bleus, et s'en attrister sont trois façons de traiter l'information que les Bleus sont éliminés, trois façons de relier cette même information à nos autres états mentaux.

Cette hybridation de la théorie attitudinale des émotions avec l'analyse des attitudes de Cummins est fort intéressante et, je crois, prometteuse. Notons toutefois qu'elle risque de trahir l'anti-représentationalisme de Deonna et Teroni, c'est-à-dire l'idée que les émotions ne représentent pas les valeurs qui font d'elles des attitudes évaluatives. Cette possible trahison théorique n'est pas un problème en soi, mais elle incite à se demander si l'hybridation de Lepine n'amène pas à une théorie attitudinale différente de l'originale plutôt qu'à un raffinement¹.

Dans ce chapitre, l'auteur propose un autre développement original de la théorie attitudinale. Celui-ci lui permet d'ailleurs de répondre à l'objection (faite notamment par Mitchell, 2021, chap. 5) selon laquelle les ressentis d'au moins certaines émotions (leur caractère phénoménal) est irréductible à leurs ressentis corporels (contrairement à la deuxième hypothèse ci-dessus). On peut penser à un émerveillement contemplatif, calme, où le corps ne semble pas jouer un rôle primordial dans l'effet que cela fait d'être dans cet état émotionnel.

La solution proposée par Lepine pour répondre à cette objection est d'insister sur la manière dont les émotions plus calmes orientent notre attention et de considérer cette orientation de l'attention comme une partie importante de notre préparation à l'action et comme un filtre par lequel passent nos ressentis corporels. Cette orientation détermine en effet quel type d'actions l'on épousera, quelles sont nos priorités dans la délibération de nos actions, et quelles modifications corporelles sont ressenties. Certes, notre corps ne subit pas de modifications radicales lors d'émotions calmes, puisqu'il ne se prépare pas à fuir ou à se battre ou à aucune autre intense réaction physique. Néanmoins, l'orientation de notre attention constitue une préparation à l'action qui joue un rôle déterminant dans le ressenti des émotions calmes. Et certains changements corporels mineurs peuvent être ressentis dans ces cas si notre attention est portée sur ceux (par exemple, le mouvement de nos yeux, notre posture, la relaxation de nos muscles). À la préparation physiologique – battements de cœur, contractions musculaires, etc. – il faut donc ajouter la préparation cognitive – orientation de l'attention – et les deux déterminent ensemble le ressenti émotionnel selon la théorie attitudinale proposée. Ce raffinement de la théorie me semble encore une fois bienvenu et prometteur (et, cette fois, respecte entièrement l'esprit de la version originale). Qui plus est, l'auteur s'appuie sur des études empiriques pour étayer cette hypothèse.

L'auteur montre dans le restant du chapitre comment sa théorie attitudinale augmentée de la composante attentionnelle permet de répondre à deux autres

objections formulées par Dokic et Lemaire (2015). Je ne vais néanmoins pas reporter ici cette discussion plus technique. Il ajoute enfin en quoi sa théorie est compatible avec la théorie psychologique de l'évaluation présentée dans les chapitres précédents, renforçant encore une fois l'unité des chapitres de ce livre. Il conclut en affirmant que la théorie attitudinale constitue à ce jour la conceptualisation la plus satisfaisante au sujet des émotions.

S'il fallait émettre une critique de ce chapitre, on pourrait reprocher l'absence de discussion de certaines théories rivales (Gordon, 1987 ; Scarantino, 2014 ; Müller, 2019 ; Mitchell, 2021) et de réponses aux objections avancées par Lepine (Tappolet, 2016 ; Rossi et Tappolet, 2019)². Néanmoins, les raffinements et ajouts de l'auteur à la théorie attitudinale – l'hybridation avec la théorie de Cummins et l'ajout de l'orientation attentionnelle – rendent ce chapitre plus que stimulant et appréciable. J'espère qu'ils susciteront les discussions qu'ils méritent.

CHAPITRE V : LES ÉMOTIONS ET LES VALEURS

Le thème de ce chapitre est le sentimentalisme : l'idée selon laquelle les concepts évaluatifs – les concepts du beau, de l'injuste, du courageux, etc. – « ne peuvent pas être compris ou analysés sans faire appel aux réponses affectives des individus, et notamment à leurs émotions et leurs motivations » (Lepine, 2023, p. 149). À ce propos, je recommande également l'entrée d'encyclopédie sur le sentimentalisme, écrite par Samuel Lepine (2018), qui complète bien ce chapitre. Ce dernier se concentre plus spécifiquement sur le versant épistémologique du sentimentalisme et la manière dont les émotions contribuent à nos croyances évaluatives. Sont aussi abordées des questions ontologiques et comment les états affectifs contribuent à la nature des valeurs morales et non-morales.

En plus de présenter ces questions, le chapitre défend plusieurs thèses originales. Notamment, l'auteur présente une version du sentimentalisme d'inspiration brentanienne selon lequel les concepts évaluatifs doivent être analysés avec la notion d'émotions correctes. Par exemple, le concept d'injustice pourrait, selon cette approche, être analysé comme ce sur quoi une indignation correcte porte. Pour analyser le concept d'injustice, il faut donc comprendre ce qu'est l'indignation – une question abordée sous sa forme générale dans les chapitres I à IV – ainsi que ce qui rend l'indignation correcte – une question abordée dans ce chapitre. Le chapitre VI viendra compléter l'analyse en abordant la question suivante : qu'est-ce qui détermine qu'une émotion est justifiée et fiable ou qu'elle ne l'est pas ?

Dans ce chapitre V, l'auteur défend plus spécifiquement une forme de sentimentalisme rationnel, selon lequel les concepts de valeurs doivent être analysés en spécifiant les raisons pertinentes qui déterminent si nos émotions sont correctes. Par « correcte », on entend ici qu'une émotion évalue correctement le monde. On a vu dans les chapitres précédents dans quel sens les émotions évaluent le monde, et on voit dans celui-ci ce qui détermine les conditions de correction de ces évaluations.

Par exemple, la tristesse impliquerait l'évaluation qu'on a subi une perte, mais qu'est-ce qui détermine si, oui ou non, on a véritablement subi une perte ?

La réponse offerte par l'auteur est une thèse qu'il qualifie de « motivationnelle » (voir Lepine, 2023, p. 162-196). Selon celle-ci, la correction des émotions dépend notamment de leur congruence avec nos motivations d'arrière-plan (les bases motivationnelles introduites au chapitre I), soit nos désirs, préférences, sentiments, ou traits de caractère contribuant à l'émotion en question. Dans l'exemple de la tristesse, le fait qu'on a subi une perte ou non dépend notamment de si l'on préférerait vraiment que l'objet de notre tristesse ne soit pas perdu. Ainsi, le fait qu'une valeur (ici la perte) soit instanciée ou non dépend des motivations de la personne concernée (de ses préférences). L'auteur défend ainsi une forme de subjectivisme au sujet de l'existence des valeurs. Cette thèse peut sembler plausible pour la perte et la tristesse, mais elle est beaucoup plus difficile à défendre pour les valeurs et émotions morales. Je reviendrai plus loin sur ce point.

L'auteur présente sa thèse motivationnelle comme une alternative à la thèse dite « indépendantiste » (défendue notamment par Tappolet (2000) mais également, plus ou moins implicitement, par Deonna et Teroni (2012), entre autres). La thèse indépendantiste veut que les conditions de correction de nos émotions soient purement objectives et indépendantes de nos motivations subjectives ou de nos jugements évaluatifs (Lepine, 2023, p. 162). En ce qui concerne le rôle que les émotions jouent vis-à-vis de nos croyances évaluatives, selon la thèse indépendantiste, les émotions peuvent être considérées comme révélant des propriétés évaluatives existant indépendamment de nos motivations (désirs, préférences, sentiments, etc.). Par contraste, la thèse motivationnelle veut plutôt que nos émotions, lorsqu'elles sont correctes, nous révèlent ce qui compte pour nous, ce qui nous préoccupe, ce qui nous motive, nous attire, nous repousse, ou nous révolte. Les émotions nous révèlent ce qui est pertinent à nos motivations d'arrière-plan ; elles ne nous révèlent pas des valeurs qui existeraient indépendamment de ces motivations.

Comme il a été dit, la thèse motivationnelle postule que les conditions des émotions dépendent des motivations qui les sous-tendent. Implique-t-elle donc que toute émotion est correcte pour autant qu'elle soit congruente avec nos préférences, désirs, sentiments, etc. ? Veut-elle dire, par exemple, que les psychopathes qui se réjouissent de torturer leurs victimes éprouvent une émotion correcte pour autant que ces psychopathes désirent accomplir ces actes ? Non, car, selon la thèse motivationnelle défendue, les émotions ne peuvent être correctes que si les motivations d'arrière-plan qui les sous-tendent sont elles-mêmes correctes (Lepine, 2023, p. 179-188).

Quels sont les critères pour déterminer si une motivation est correcte ? Les motivations pertinentes (désirs, préférences, sentiments, traits de caractère) peuvent être considérées comme correctes, selon Lepine, aussi longtemps qu'on n'a pas de raison de croire qu'elles devraient être rejetées ; lorsqu'il n'y a pas de « défaites » (il s'inspire ici librement du faillibilisme de Tappolet [2000] sans l'appliquer aux mêmes fins) : « Ainsi, le critère de correction proposé ici revient à

dire qu'une motivation est correcte aussi longtemps que nous ne connaissons pas de défauteurs qui s'opposeraient à elle. » (Lepine, 2023, p. 180)

Comme exemple de motivations incorrectes, l'auteur mentionne notamment que certaines phobies et idéologies racistes sont des motivations qui reposent en moins en partie sur des croyances injustifiées et que ces dernières sont des défauteurs rendant ces motivations incorrectes. Les émotions des phobiques et racistes en question sont donc, à leur tour, incorrectes. Il donne également l'exemple de considérations normatives, notamment morales, qui pourraient jouer le rôle de défauteurs en nous faisant réaliser que des motivations d'arrière-plan doivent être rejetées : elles seraient incorrectes en cela que les « meilleures raisons disponibles dans notre environnement » (Lepine, 2023, p. 180) sont des défauteurs des motivations en question et nous amènent (ou devraient nous amener) à les rejeter.

On pourrait donc imaginer que les psychopathes qui se réjouissent de leurs atrocités éprouvent une émotion incorrecte en cela que les meilleures raisons disponibles dans leur environnement, des raisons qu'ils et elles connaissent, devraient les amener à rejeter le sadisme et autres motivations qui fondent leur cruelle réjouissance. Mais pourquoi ne pas penser (et c'est sans doute ce que les psychopathes en question se disent) que le plaisir tiré de leurs activités constitue une raison pour leur sadisme qui, tout bien considéré, est la meilleure disponible dans cet environnement ? Il semble que l'auteur répondrait à cet exemple de la façon suivante : il faut distinguer différentes conditions de correction pour la joie des psychopathes. Certes, il est possible que les psychopathes en question n'aient pas de raisons *prudentielles* (qui concernent ce qui est bien pour elles et eux) disponibles dans leur environnement qui les amèneraient à rejeter leur motivation sadique, mais il existe certainement des raisons morales auxquelles ils et elles auraient accès et qui devraient les amener à rejeter ces motivations-là. Ainsi, leur joie n'est peut-être pas incorrecte d'un point de vue prudentiel, mais elle l'est d'un point de vue moral (cf. les remarques de l'auteur sur l'hilarité ressentie envers une blague immorale, p. 189). Pour étayer cette réponse, il faudra toutefois expliciter en quoi les psychopathes en question ont forcément accès à des raisons morales qui devraient les amener à rejeter leur motivation sadique. Ou accepter que leur joie est moralement correcte. Ou modifier le critère de correction des états motivationnels. Je reviendrai sur ces questions.

Prenons un autre exemple, mentionné par Lepine : un enfant élevé dans un environnement raciste, rempli de nazis et de membres du Ku Klux Klan, qui n'aura pas accès – étant donné son expérience passée, ses sources d'information, et ses capacités cognitives – à des considérations qui seraient des défauteurs des sentiments racistes qu'on lui a inculqués. Il admet que l'environnement de cet enfant lui permet d'avoir des croyances racistes cohérentes et qu'il n'aurait pas de raisons de rejeter ses sentiments racistes – au contraire : les meilleures raisons à disposition dans son environnement soutiennent son racisme. Il n'y a pas de défauteurs pour ses motivations racistes. Lepine semble ainsi admettre que les émotions racistes de cet enfant seraient, selon la thèse défendue, correctes (bien

que cela ne soit pas sûr étant donné les deux derniers paragraphes du chapitre – j’y reviendrai). Ce contre-exemple apparent ne semble néanmoins pas l’inquiéter.

Il écrit à son propos :

Cette objection, en réalité, n’est pas une grande menace pour la thèse motivationnelle. La plupart du temps, notre expérience nous donne les moyens de distinguer nos mauvaises motivations. Un raciste sait pertinemment que tous les individus qu’il déteste ne correspondent pas à l’idée *a priori* qu’il s’en fait ou qu’on lui a inculquée. (Lepine, 2023, p. 182)

Je dois dire que je suis moins optimiste quant aux croyances des racistes. Bien sûr, le cas de l’enfant élevé uniquement par des nazis et membres du KKK est extraordinaire. Mais, me semble-t-il, il n’est pas vrai que les racistes savent pertinemment que les idées qu’ils ou elles se font des personnes détestées sont fausses. L’hypothèse que, la plupart du temps, notre expérience nous donne les moyens de distinguer nos mauvaises motivations ne serait vraie, selon moi, que si nous étions des êtres épistémiquement idéaux ou du moins si on s’en rapprochait. Mais nos biais innés et acquis, nos limites cognitives, notre incroyable capacité à confabuler – c’est-à-dire à combler les lacunes de la mémoire par des inventions souvent arrangeantes –, ainsi que les limites structurelles de notre environnement (reproduction sociale, chambres d’écho, propagande, inaccessibilité des informations) me laissent penser que, non, la plupart du temps, nous n’avons pas les moyens de distinguer nos mauvaises motivations. Il me semble que nos limites cognitives et nos environnements nous empêchent souvent de savoir que nos mauvaises motivations sont basées sur des évaluations ou des croyances injustifiées.

À mon sens, c’est là un problème potentiel de l’approche proposée par Lepine car celle-ci semble, au moins dans certains passages, tendre vers une forme de faillibilisme cohérentiste qui voudrait que, tant que les motivations et les croyances d’une personne sont cohérentes entre elles, tout état motivationnel de cette personne est correct puisque, dans une telle situation de cohérence, « notre expérience nous donne des raisons de penser qu’il [l’état motivationnel] est justifié, et qu’il n’a pas été pris en défaut par de meilleures raisons » (Lepine, 2023, p. 181). Ainsi, selon cette interprétation, tant que les motivations d’une personne sont cohérentes avec ses croyances, les émotions qui sont congruentes avec ses motivations sont à leur tour correctes.

Cela pose problème car il est difficile d’accepter que l’enfant raciste mentionné ait des émotions racistes correctes (même si on les considère comme justifiées en l’absence de défauteurs). Or, comme l’a bien souligné Williamson (2019), un cas de racisme cohérent n’est pas si extraordinaire que Lepine le laisse suggérer. Selon Williamson, il est en fait probable qu’on puisse être un néonazi dont les croyances et les motivations sont cohérentes entre elles, surtout lorsqu’on considère que nos biais (comme le biais de confirmation) et nos capacités à

confabuler nous poussent à croire des choses cohérentes avec nos motivations, même lorsque ces croyances semblent, d'un point de vue externe, évidemment fausses. Le néonazi de Williamson est cohérent car il rejette, de par ses biais et ses confabulations, les croyances qui s'opposent à son idéologie : dans un sens en tout cas, il ne connaît pas (il n'a pas conscience) de raisons qui sont des défauteurs de ses motivations racistes. Pour lui, « il n'existe aucune considération, à [sa] connaissance, qui pourrait militer contre elle[s] » (Lepine, 2023, p. 181), dans le sens où il n'a pas un accès conscient à de telles considérations. Si « connaître » des raisons implique, dans la théorie de Lepine, « avoir un accès conscient » à ces raisons, alors les motivations racistes de ce néonazi cohérent et les émotions qui en découlent seraient correctes, selon l'interprétation cohérentiste du faillibilisme de Lepine.

La question de savoir si les personnes dotées de motivations racistes, sadiques, sexistes, homophobes, ou d'autres motivations vicieuses sont souvent ou non cohérentes est, *in fine*, une question empirique. Peut-être l'auteur a-t-il raison de croire que, la plupart du temps, ces personnes savent que leurs motivations devraient être rejetées – qu'elles sont « défaites » par les meilleures raisons à leur disposition. Mais la possibilité même que ce ne soit pas le cas me semble être une menace sérieuse pour la thèse motivationnelle dans son interprétation cohérentiste. Il suffit qu'il y ait une personne dans un monde possible qui ne connaisse pas de raisons qui sont des défauteurs pour ses motivations vicieuses (racistes, sadiques, etc.) pour soulever la présente objection : il n'est pas vrai que les émotions découlant de ces motivations vicieuses sont correctes – ni la joie sadique, ni le dégoût raciste, ni les émotions sexistes, homophobes, etc. Ainsi, semble-t-il, la lecture cohérentiste de la correctitude contraste fortement avec ce que l'on entend généralement lorsque l'on dit qu'une émotion est correcte.

Que répondrait l'auteur à cette objection ? Il est possible qu'il rejette la lecture purement cohérentiste que je fais des conditions de correction des états motivationnels (Lepine, 2023, p. 180 et 181) – peut-être rejetterait-il l'idée que « connaître » des raisons qui sont des défauteurs impliquerait « en avoir conscience ». Ainsi, à certains endroits, il semble suggérer que, dans les exemples cités, il doive y avoir des raisons morales qui sont accessibles aux personnes en question : selon cette réponse, contrairement à ce que j'ai supposé, l'enfant élevé par les racistes et le néonazi cohérent de Williamson auraient forcément un accès à des raisons morales qui seraient des défauteurs pour leurs motivations racistes. De même, les psychopathes dont je discutais auraient forcément accès à des raisons morales qui rendraient leurs motivations incorrectes. Toutefois, cette réponse devrait nous dire pourquoi il n'existe aucun monde possible avec une personne qui, d'une part, a des motivations vicieuses dont découlent des émotions moralement incorrectes (racistes, sadiques) et qui, d'autre part, n'a pas accès à des raisons morales qui sont des défauteurs de ces motivations.

Passons à un autre sujet : l'ontologie des valeurs proposée par Lepine. Comme mentionné plus haut, sa thèse motivationnelle l'amène à défendre une forme de subjectivisme quant à l'existence des valeurs et en particulier la thèse selon

laquelle l'existence de valeurs dépend de nos motivations. On a vu que, selon cette thèse, le fait qu'on a ou non subi une perte dépend de si l'on préférerait vraiment que l'objet de notre tristesse ne fût pas perdu. Mais peut-on étendre cette idée à toute valeur et toute émotion ? Peut-on vraiment dire que ce qui fait qu'une indignation est correcte ou non dépend des préférences de la personne indignée ? Cela ne reviendrait-il pas à dire qu'il n'y a une injustice que lorsqu'il y a une préférence pour l'injustice ? L'auteur laisse la porte ouverte à une réponse positive et à une forme de relativisme ontologique vis-à-vis des valeurs, y compris morales (il cite notamment le relativisme moral de Prinz [2007] comme modèle potentiel) :

D'un point de vue ontologique, en effet, l'approche motivationnelle est subjectiviste, au sens où elle soutient que les valeurs désignent ce que nous sommes disposés à apprécier en fonction de nos motivations subjectives, et en l'absence de tout défaiteur. [...] Cette approche implique donc une ontologie des valeurs très largement relativiste, puisque les valeurs ici sont toujours relatives à nos motivations. (Lepine, 2023, p. 193-194)

Ainsi, semble-t-il, une injustice n'existe qu'en vertu de l'indignation qu'elle provoquerait. Mais cette théorie semble impliquer – comme le suggèrent Deonna et Teroni (2012, chap. 4) – que, dans le cas où l'humanité serait réduite à un état de barbarie totale à la suite d'une guerre apocalyptique et que, un peu comme dans *Mad Max* ou *Escape from L.A.*, tout le monde en venait à considérer l'esclavage, le pillage, ou le viol comme des actes aussi banals qu'un brossage de dents, ce monde serait plus juste que le nôtre. Dans une telle situation, la cruauté serait devenue la norme et la société serait basée sur un rejet de préoccupation d'équité. L'indignation en serait pratiquement absente de même que les motivations la sous-tendant. Mais, par conséquent, selon le relativisme proposé, l'injustice serait pratiquement absente, puisqu'il serait si rare qu'on soit disposé à être indigné. Le nombre de choses injustes serait donc pratiquement égal à zéro ; ce monde serait plus juste que le nôtre.

L'auteur est-il forcé d'avalier cette couleuvre ? Non, car il ouvre la porte, dans l'avant-dernier paragraphe du chapitre V à l'idée qu'il existe un « petit nombre de vérités morales élémentaires [...] des règles que tout le monde pourrait accepter » (Lepine, 2023, p. 196) qui seraient fondées sur notre aversion pour la souffrance ou la douleur. Il cite Singer (1979) comme modèle potentiel (un modèle autre que le relativisme de Prinz, donc). Lepine admet que cette option éloignerait son ontologie du relativisme moral et pourrait le conduire à la place vers une forme d'objectivisme dans ce domaine. Peut-être cette piste pourrait-elle lui permettre de répondre à l'autre objection soulevée ci-dessus impliquant les racistes et les psychopathes ? L'auteur ne développe pas davantage cette question. Peut-être le fera-t-il dans de futurs ouvrages ?

D'autres conséquences de la thèse motivationnelle sont discutées dans ce chapitre et l'auteur répond aux problèmes potentiels qui en découlent. Je me contenterai de les énumérer :

- Comment la thèse motivationnelle permet-elle de rendre compte qu'il existe de véritables découvertes émotionnelles et de l'intuition que les émotions semblent parfois nous révéler des propriétés évaluatives jusque-là inconnues ?
- La thèse motivationnelle est-elle une forme de pragmatisme au sujet des valeurs ? L'auteur défend que non et soutient une version « épistémique » plutôt que pragmatique de la thèse motivationnelle.
- Comment la thèse motivationnelle peut-elle résoudre les problèmes de la confusion émotionnelle et des mauvais types de raisons ?

J'ai insisté sur les problèmes potentiels liés à un faillibilisme cohérentiste et au relativisme moral car ce sont, selon moi, les menaces les plus pressantes pour la thèse motivationnelle de l'auteur et parce que les solutions suggérées me semblent mériter des discussions plus approfondies. Cela est peut-être injuste de ma part car ce sont sans doute des détails dans l'ensemble de ce chapitre qui présentent des arguments sophistiqués et, selon moi, très satisfaisants, notamment sur les trois questions que j'ai seulement rapportées ici. J'ajouterai encore que l'ontologie des valeurs naturaliste défendue par l'auteur me semble, dans l'ensemble, fort attrayante et, malgré les problèmes potentiels que j'ai soulevés, elle fait de sa thèse motivationnelle un rival plus que sérieux à la thèse indépendantiste de Tappolet ou Deonna et Teroni en ce qui concerne les conditions de correction des émotions.

CHAPITRE VI : LA JUSTIFICATION ET LA FIABILITÉ DES ÉMOTIONS

Dans le sixième et dernier chapitre sont abordées deux questions épistémiques qu'on pourrait considérer comme des défis à la thèse motivationnelle défendue dans le chapitre précédent. La première question concerne la relation entre la justification des émotions et les motivations d'arrière-plan qui expliquent les émotions : peut-on vraiment soutenir, comme la thèse motivationnelle l'implique, que l'amour de Michelle pour sa fille justifie sa joie de la voir (Lepine, 2023, p. 197) ? Quid du cas des psychopathes discuté ci-dessus : leur sadisme justifie-t-il leur joie de causer la souffrance d'autrui ? La deuxième question du chapitre concerne la fiabilité des émotions dans leur capacité, supposée par la thèse motivationnelle, à nous faire appréhender les valeurs : les émotions ne biaisent-elles pas nos jugements de valeur ? Ne nous poussent-elles pas à confabuler et inventer des justifications *ad hoc* qui nous confortent dans nos préjugés ?

En réponse à la première question, l'auteur maintient que les motivations d'arrière-plan justifient les émotions qu'elles déclenchent – elles en sont à la fois des raisons explicatives (étant l'une de leurs causes) et des raisons justificatives. Par exemple, le fait que Sam se préoccupe de son poids pour des raisons de santé explique, d'une part, qu'il ressente de la joie lorsqu'il réalise qu'il a perdu ses kilos en trop mais, d'autre part, cette préoccupation justifie également sa joie : c'est notamment parce qu'il souhaite être en bonne santé qu'il a une bonne raison de se réjouir. Néanmoins, l'auteur ajoute une condition importante, déjà évoquée

au chapitre précédent : les motivations d'arrière-plan ne justifient les émotions que lorsque ces motivations sont elles-mêmes justifiées. Nous pouvons avancer des raisons d'endosser nos motivations ou de les abandonner, et cela va, à son tour, permettre de justifier les émotions qui en dérivent. Les motivations sont ainsi l'objet de « justification primaire » et les émotions de « justification secondaire ». Plus précisément, l'auteur avance que certains types de motivations sont susceptibles de servir de raisons primaires en raison de leurs liens intimes avec ce qui a de la valeur pour la personne qui possède ces motivations :

- Les sentiments, compris comme des attachements à, ou des aversions envers, des choses, personnes, animaux, ou institutions. Par exemple, être francophile, c'est avoir un attachement à ce qui est français.
- Les traits de caractère, compris comme des sensibilités à certaines valeurs. Par exemple, être bienveillant serait être particulièrement sensible au bien d'autrui.
- Les désirs, compris comme des états qui représentent leurs objets comme « devant être » (l'auteur suit ici l'analyse des désirs de Lauria, 2014). Désirer manger une glace impliquerait de voir cet état de choses comme devant être le cas.
- Enfin, les préférences, comprises comme les représentations de valeurs subjectives, des états ayant pour fonction d'indiquer ce qui est bon pour soi. Avoir une préférence pour passer du temps avec nos proches, ce serait représenter cela comme bon pour nous.

Les préférences sont d'ailleurs considérées par l'auteur comme les motivations les plus fondamentales, celles sur lesquelles sont indexés tous les autres types de motivations (Lepine, 2023, p. 205). Ainsi, les préférences sont, pour l'auteur, l'ultime source évaluative de la justification des jugements de valeur : on peut justifier qu'on trouve bon, beau, bête quelque chose par une émotion ressentie, laquelle peut être justifiée par nos sentiments, traits de caractère ou désirs, et ceux-là sont indexés sur nos préférences, lesquelles doivent être justifiées : représentent-elles correctement ce qui est bon pour nous ? Sont-elles vraiment dans notre intérêt objectif ?

L'auteur n'offre pas une théorie des intérêts objectifs, c'est-à-dire de ce qui est réellement bon pour soi. Toutefois, peut-être que c'est là qu'il faudrait trouver la réponse à une question que j'ai posée ci-dessus quant à la justification de la joie des psychopathes. Si l'on part du principe que leur joie est injustifiée, alors il faut conclure, semble-t-il, qu'il ne serait, en fait, pas dans leurs intérêts – il ne serait en fait pas bon pour ces psychopathes – de faire souffrir les autres. Leurs préférences sadiques échouent dans leurs fonctions représentationnelles. Leur joie ne sera donc pas justifiée par leur sadisme.

Cette hypothèse invite à se demander pourquoi il ne serait pas bon pour les psychopathes que les autres souffrent de leurs mains. Ce serait peut-être pour des raisons prudentielles – cela les empêcherait d'atteindre un bien-être durable, leur interdisant de profiter du bonheur que leur apporteraient leurs pairs. Ou alors

pour des raisons morales – on reviendrait peut-être à ces « vérités élémentaires » de la morale mentionnées au chapitre précédent pour expliquer quels sont les véritables intérêts des psychopathes en question. Évidemment, l’auteur pourrait également refuser le point de départ de mon raisonnement et soutenir que la joie des psychopathes en question est, en fait, justifiée – prudemment, cela paraît défendable, mais moralement, c’est difficile de voir comment. Si mon raisonnement est correct, la théorie de Lepine et les intuitions que j’ai mises en avant semblent mener à l’idée que la joie des psychopathes serait incorrecte, *in fine*, car leur sadisme est indexé sur une préférence qui représente incorrectement, pour des raisons morales, ce qui est bon pour eux (et cela en vertu du fait que la meilleure raison morale à leur disposition est un défauteur pour cette préférence). Selon moi, un développement de ces points serait un complément fort bienvenu à ce livre qui est déjà par ailleurs très riche.

Venons-en à la deuxième question abordée dans ce chapitre : la fiabilité des émotions. Sur ce point, l’auteur s’efforce de défendre l’idée que les émotions – tout comme les motivations – « sont des états suffisamment fiables pour que nous puissions leur faire confiance le plus souvent » (Lepine, 2023, p. 224). La stratégie employée est de démontrer que les deux objections principales à cette idée sont en fait sans fondement.

La première des deux objections veut que les émotions ne soient pas fiables puisque ce sont des attitudes psychologiques qui ont évolué dans des contextes différents et, bien qu’elles aient été adaptées à la niche écologique de nos ancêtres distants, elles sont souvent inadéquates dans l’environnement actuel de la plupart des êtres humains. Lepine cite Goldie (2008) qui défend qu’au moins trois types d’émotions ne sont plus adaptées pour ces raisons :

- La fierté et la honte relatives aux agressions entre mâles.
- La peur et la méfiance vis-à-vis des étrangers.
- La jalousie sexuelle et les violences qui l’accompagnent.

Lepine répond à cette objection en notant, d’abord, que ces cas-là devraient plutôt être analysés comme étant primordialement des motivations injustifiées plutôt que des émotions inadéquates. Il ajoute, d’autre part, que les motivations problématiques en question peuvent être rejetées par les personnes qui les ont – celles-ci ne sont pas condamnées à vivre avec elles. Nos motivations et nos émotions sont plastiques, et nous pouvons donc nous débarrasser d’une fierté, honte, peur, méfiance ou jalousie injustifiée.

Notons que cette double réponse n’empêche pas de penser que ces motivations inadéquates sont néanmoins répandues et rendent ainsi nombre de nos émotions injustifiées et inadéquates. Après tout, la xénophobie et les violences en question semblent être malheureusement bien trop communes. La réponse ne montre donc pas que nos émotions et nos motivations sont des états suffisamment fiables pour qu’on puisse leur faire confiance la plupart du temps.

La discussion de la deuxième objection considérée par l’auteur permet en partie de répondre à cette inquiétude (mais pas de façon entièrement satisfaisante, à mon sens). L’objection en question veut que les émotions ne soient pas fiables parce qu’elles ont tendance à engendrer des justifications qui biaisent nos jugements et ainsi poussent à des confabulations qui nous empêchent de réaliser que nos émotions et nos motivations sont inadaptées (Lepine, 2023, p. 211). Par exemple, lorsqu’on se met en colère, on a tendance à porter notre attention sur des détails qui nous permettent de justifier notre colère et à ignorer ceux qui ne vont pas dans notre sens – le biais de confirmation serait ainsi exacerbé par les émotions.

L’auteur développe de façon remarquable ce point en consacrant plusieurs pages à expliciter la façon dont nous avons tendance à confabuler et à nous duper nous-mêmes – le tout agrémenté de plusieurs études empiriques sur la question ainsi que de travaux philosophiques sur la relation de ces biais aux émotions (notamment Brady, 2013). Il conclut de cette discussion qu’un regard critique doit être exercé sur nos émotions et sur les conclusions auxquelles peuvent nous amener celles-ci, mais que, heureusement, ce rôle est déjà joué en bonne partie par notre système affectif lui-même, puisque nous sommes conduits à exercer ce regard critique dès lors que nos émotions entrent en conflit avec certaines de nos motivations (un sujet discuté au chapitre précédent). Une sorte de vigilance épistémique nous protégerait donc des biais liés aux émotions en cela que, dès lors qu’un manque de cohérence entre nos émotions et nos motivations émerge, il est détecté – nous ressentons ce manque de cohérence. Cela nous permet ainsi de corriger nos émotions et motivations inadéquates.

Certes, l’auteur reconnaît que, étant donné notre propension à confabuler et notre biais de confirmation (entre autres biais cognitifs), lorsque nos motivations sont inadéquates, cela risque de nous pousser à être cohérents avec ces motivations inadéquates plutôt qu’avec les données probantes auxquelles nous avons eu accès. Nous rejeterions donc parfois des émotions incohérentes avec nos motivations en les traitant comme étant injustifiées alors qu’en fait elles sont parfaitement adéquates. On peut par exemple imaginer des xénophobes qui se disent « Non! Ces réfugiés ne méritent pas ma compassion. Quelle faiblesse de ma part d’avoir éprouvé quoi que ce soit pour cette vermine. »

Lepine conclut donc qu’il faut se méfier de nos motivations d’arrière-plan bien plus que des émotions qui en découlent car ce sont les motivations qui rendent certaines émotions inadaptées et nous poussent à confabuler. Et lorsque nos motivations sont adaptées, les émotions qui en découlent sont certes partiales – elles concentrent notre attention sur certains détails en nous en faisant ignorer d’autres – mais elles le sont de façon vertueuse, en nous aidant à nous focaliser sur – voire à découvrir – ce qui compte vraiment pour nous, ce à quoi nous tenons le plus, et le genre de personne que nous voulons véritablement devenir (Lepine, 2023, p. 224).

J'adhère à cette conclusion, soit que nos motivations d'arrière-plan, et en particulier nos préférences les plus basiques, sont ce à quoi nous devons porter le plus d'attention, car c'est seulement des motivations adaptées que découlent les émotions adaptées (cf. Bonard, 2021a, chap. 9 ; 2024, où les motivations de base sont appelées « buts » (*goals*) plutôt que préférences). Néanmoins, la discussion sur la fiabilité des émotions de Lepine me semble prêter le flanc à deux difficultés principales.

La première a déjà été soulevée vis-à-vis du chapitre précédent : étant donné l'interprétation cohérentiste que j'ai proposée des conditions de correction des états motivationnels, celles-ci ne permettent pas de rendre compte des cas où une personne – comme le néonazi de Williamson – est cohérente sur le long terme alors que ses motivations sont inadaptées. Cette personne pourra, certes, détecter des émotions ou des croyances incohérentes avec ses motivations grâce à la supposée vigilance inhérente à son système affectif, mais elle les rejettera, quitte à se forger une vie mentale largement confabulée. Le fait que nous possédions sans doute des motivations inadaptées héritées de l'évolution rend cette inquiétude d'autant plus pressante – on peut penser aux plausibles exemples de Goldie mentionnés plus haut.

Notons aussi que la plasticité des émotions et des motivations n'aide pas forcément dans ces cas-là. Au contraire, cette plasticité implique que l'on peut rejeter des émotions incohérentes avec nos motivations alors que ce sont ces dernières qui sont inadéquates. Les xénophobes que je mentionnais en exemple pourraient très bien se débarrasser de leur compassion envers les réfugiés en vertu de la plasticité des émotions et motivations.

La deuxième difficulté majeure que je vois quant à la fiabilité des émotions est qu'il y a d'autres raisons que celles discutées par l'auteur de penser que les émotions manquent systématiquement de fiabilité. Je pense en particulier au cas suivant : comme l'a souligné avec force Godfrey-Smith (1991), certains mécanismes que nous a laissés l'évolution sont censés se déclencher trop souvent. Par exemple, il est normal, d'un point de vue évolutionnaire, que la peur se déclenche plus souvent dans des situations où nous ne sommes pas en danger que dans celles où nous le sommes, car le coût de ces déclenchements excessifs est bien moindre que le coût de ne pas réagir avec peur à des dangers réels. Mieux vaut qu'une alarme incendie se déclenche trop souvent, alors qu'il n'y a pas de feu, que pas assez souvent, alors qu'il y en a. Pour vivre plus longtemps, mieux vaut qu'on se trompe régulièrement en croyant détecter un danger plutôt que de ne pas se méfier de ce qui peut causer sa perte. Comme le dit Goffin (2023) en analysant ce phénomène : *better be scared than sorry*.

Ainsi, il est probable que la peur, par nature, ne soit pas fiable et se déclenche plus souvent quand il n'y a pas de danger que lorsqu'il y en a un. Cette hypothèse me semble par ailleurs être corroborée par notre expérience quotidienne. Et remarquons que cela reste vrai même si l'on suppose que les motivations d'arrière-plan sont correctes et justifiées – par exemple, que le désir de rester sain

et sauf est approprié. Il est par ailleurs plausible que ces remarques s'étendent à d'autres émotions que la peur – comme le dégoût, la colère, la jalousie, la honte, ou l'espoir – pour des raisons évolutives semblables

CONCLUSION

Bien que je me sois efforcé dans une bonne partie de cette recension de faire des remarques critiques et pointer du doigt des questions qui me semblaient mériter une plus ample discussion, il ne faut pas en conclure que mon jugement général est négatif. Bien au contraire, les questions et les critiques que j'ai formulées sont à considérer avant tout comme des invitations, adressées aux lecteurs et lectrices, à discuter plus en profondeur les idées originales et stimulantes de Samuel Lepine.

J'ai trouvé ce livre excellent et je considère que c'est à ce jour la meilleure introduction à la philosophie des émotions en français. Mais le livre va bien au-delà d'une introduction : il présente toute une série d'idées originales, plausibles, stimulantes et bien argumentées. Convaincu par de nombreuses thèses avancées par l'auteur, je suis globalement satisfait de la façon dont les discussions de philosophie et de psychologie des émotions ont été présentées. J'ai également appris énormément grâce à l'impressionnante richesse de la littérature abordée, et la diversité des sources discutées a soulevé mon enthousiasme. J'ai ressenti énormément d'émotions positives à la lecture de ce livre – en vertu de motivations correctes, je l'espère.

Les six chapitres de ce livre forment un ouvrage cohérent, riche, solide, original et édifiant. Le tout peint un portrait convaincant de ce à quoi un sentimentalisme contemporain peut ressembler, un sentimentalisme basé sur les découvertes des sciences empiriques et les ouvrages philosophiques les plus pertinents de ces dernières décennies.

Le souhait de l'auteur, exprimé dans l'introduction, était que ce livre puisse être lu par différents publics : des néophytes aux spécialistes en passant par les étudiants et étudiantes en philosophie. Le pari me semble tout à fait réussi et je souhaite à mon tour que cette recension contribue un peu à ce que ce souhait se réalise.

NOTES

¹ Ainsi, on a presque envie de résumer l'hybridation en disant que les types d'émotions sont des attitudes qui ont chacun leur propre *mode de présentation* de l'information sur lesquelles les émotions portent, bien que l'auteur évite cette expression frégeenne et parle de « sens » ou de « fonction » et que Cummins parle de « contenu sémantique » (trois expressions néanmoins très proches de celle de « mode de présentation »). Or, Deonna et Teroni s'opposent (2015) explicitement à une lecture frégeenne des types d'émotions comme ayant chacun leur propre type de mode de présentation. Cette dernière idée reviendrait à dire que les émotions *représentent* des valeurs au sein de leur mode de présentation – c'est-à-dire au sein de leur contenu « complet » au sens de Recanati (2007, pp. 20–23) lequel inclue références et modes de présentation.

² Pour les initié·es, j'ajouterai à ce propos que de nombreux travaux récents sur les croyances (voir la recension de Porot et Mandelbaum, 2021) ainsi que sur la nature des contenus propositionnels (par exemple, Mitchell, 2019a), permettent de défendre une forme des théories judicative ou mixte contre l'objection selon laquelle les états propositionnels ne sont pas accessibles aux bébés ou aux animaux non-humains (je défends moi-même une théorie mixte dans un article en cours de rédaction). Pour une discussion pertinente à ce sujet parue après que Lepine ait fini son livre, je recommande l'excellent article de Pendoley (2023).

BIBLIOGRAPHIE

- Bonard, Constant, *Meaning and Emotion: The Extended Gricean Model and What Emotional Signs Mean*, these de doctorat, University of Geneva et University of Antwerp, 2021a.
- , « Émotions et sensibilité aux valeurs : quatre conceptions philosophiques contemporaines », *Revue de métaphysique et de morale*, vol. 110, no. 2, 2021b, p. 209-229.
- , « The Rationality of Mood », in Julien Deonna, Christine Tappolet et Fabrice Teroni (dir.), *A Tribute to Ronald de Sousa*, 2022, <https://www.unige.ch/cisa/related-sites/ronald-de-sousa/>.
- , « Emotions Represent Evaluative Properties Unconsciously », *Erkenntnis*, 2024.
- Brady, Michael S., *Emotional Insight: The Epistemic Role of Emotional Experience*, Oxford, Oxford University Press, 2013.
- Cummins, Robert, *Representations, Targets, and Attitudes*, Cambridge, MIT Press, 1996.
- Deonna, Julien A. et Fabrice Teroni, *The Emotions: A Philosophical Introduction*, Londres, Routledge, 2012.
- , « Emotions As Attitudes », *dialectica*, vol. 69, no. 3, 2015, p. 293-311.
- Dokic, Jérôme et Stéphane Lemaire, « Are Emotions Evaluative Modes? », *dialectica*, vol. 69, no. 3, 2015, p. 271-292.
- Fodor, Jerry, *The Modularity of Mind*, Cambridge, MIT Press, 1983.
- Godfrey-Smith, Peter, « Signal, Decision, Action », *The Journal of Philosophy*, vol. 88, no. 12, 1991, p. 709-722.
- Goffin, Kris, « Better Scared Than Sorry: The Pragmatic Account of Emotional Representation », *Erkenntnis*, vol. 88, no. 6, 2023, p. 2633-2650.
- Goldie, Peter, « Misleading Emotions », in Georg Brun, Ulvi Doğuoglu et Dominique Kuenzle (dir.), *Epistemology and Emotions*, Londres, Routledge, 2008, p. 149-165.
- Gordon, Robert M., *The Structure of Emotions: Investigations in Cognitive Philosophy*, Cambridge, Cambridge University Press, 1987.
- Griffiths, Paul, *What Emotions Really Are: The Problem of Psychological Categories*, Chicago, University of Chicago Press, 2008.
- Lauria, Federico, *The Logic of the Liver: A Deontic View of the Intentionality of Desire*, thèse de doctorat, University of Geneva, 2014.
- Lepine, Samuel, « Sentimentalisme (A) », *L'Encyclopédie philosophique*, 2018, <http://encyclophilo.fr/sentimentalisme-a/>.
- , *La nature des émotions: une introduction partisane*, Paris, Vrin, 2023.
- Mitchell, Jonathan, « Emotional Experience and Propositional Content », *dialectica*, vol. 73, no. 4, 2019a, p. 535-561.
- , « The Intentionality and Intelligibility of Moods », *European Journal of Philosophy*, vol. 27, no. 1, 2019b, p. 118-135.
- , *Emotion as Feeling Towards Value: A Theory of Emotional Experience*, Oxford, Oxford University Press, 2021.
- Moors, Agnes, *Demystifying Emotions: A Typology of Theories in Psychology and Philosophy*, Cambridge, Cambridge University Press, 2022.
- Müller, Jean Moritz, *The World-Directedness of Emotional Feeling: On Affect and Intentionality*, Cham, Palgrave-Macmillan, 2019.
- Mulligan, Kevin, « On Being Struck by Value – Exclamations, Motivations and Vocations », in Barbara Marker (dir.), *Leben mit Gefühlen: Emotionen, Werte und ihre Kritik*, Paderborn, Mentis, 2009, p. 139-161.

- Pendoley, Kate, « Stubborn Emotions, Stubborn Beliefs », *Synthese*, vol. 201, no. 5, 2023.
- Porot, Nicolas et Eric Mandelbaum, « The Science of Belief: A Progress Report », *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 12, no. 2, 2021, p. e1539.
- Prinz, Jesse, « Are Emotions Feelings? », *Journal of Consciousness Studies*, vol. 12, nos 8-9, 2005, p. 9-25.
- , *The Emotional Construction of Morals*, Oxford, Oxford University Press, 2007.
- Recanati, François, *Perspectival Thought: A Plea for Moderate Relativism*. Oxford, Oxford University Press, 2007.
- Rossi, Mauro, « A Perceptual Theory of Moods », *Synthese*, vol. 198, no. 8, 2021, p. 7119-7147.
- Rossi, Mauro et Christine Tappolet, « What Kind of Evaluative States Are Emotions? The Attitudinal Theory vs. the Perceptual Theory of Emotions », *Canadian Journal of Philosophy*, vol. 49, no. 4, 2019, p. 544-563.
- Scarantino, Andrea, « The Motivational Theory of Emotions », in Justin D'Arms et Daniel Jacobson (dir.), *Moral Psychology and Human Agency*, Oxford, Oxford University Press, 2014, p. 156-185.
- Singer, Peter, *Practical Ethics*, Cambridge, Cambridge University Press, 1979.
- Smith, Ryan et Richard D. Lane, « Unconscious Emotion: A Cognitive Neuroscientific Perspective », *Neuroscience & Biobehavioral Reviews*, vol. 69, 2016, p. 216-238.
- Sperber, Dan, *La contagion des idées*, Paris, Odile Jacob, 1996.
- Tappolet, Christine, *Émotions et valeurs*, Paris, Presses Universitaires de France, 2000.
- , *Emotions, Value, and Agency*, Oxford, Oxford University Press, 2016.
- Whiting, Demian, « The Feeling Theory of Emotion and the Object-Directed Emotions », *European Journal of Philosophy*, vol. 19, no. 2, 2011, p. 281-303.
- Williamson, Timothy, « Morally Loaded Cases in Philosophy », *Proceedings and Addresses of the American Philosophical Association*, vol. 93, 2019, p. 159-172.
- Winkielman, Piotr et Kent C. Berridge, « Unconscious Emotion », *Current Directions in Psychological Science*, vol. 13, no. 3, 2004, p. 120-123.