

Université de Montréal

Revue Critique des Arguments en Faveur et à l'Encontre  
de la Théorie Humienne de la Motivation

Patrick Soulières

Département de Philosophie  
Faculté des Arts et des Sciences

Mémoire présenté à la Faculté des Arts et des  
Sciences en vue de l'obtention du grade de  
M.A. en Philosophie (option Recherche)

Décembre 2009

© Patrick Soulières 2009

Université de Montréal  
Faculté des Arts et des Sciences

Ce mémoire intitulé :

**Revue Critique des Arguments en Faveur et à  
l'Encontre de la Théorie Humienne de la Motivation**

présenté par :

**Patrick Soulières**

a été évalué par un jury composé des personnes suivantes :

Ryoa Chung

.....  
président-rapporteur

Christine Tappolet

.....  
directrice de recherche

Daniel Laurier

.....  
membre du jury

## **RÉSUMÉ**

Bien qu'elle constitue toujours l'orthodoxie en philosophie de l'esprit, la théorie humienne de la motivation (selon laquelle la motivation origine toujours d'un désir et jamais d'une croyance seule) a été plusieurs fois critiquée au cours des dernières décennies. Je passe ici en revue les principaux arguments avancés à la fois pour appuyer et pour rejeter cette théorie. Du côté humien, j'examine l'argument des croyances identiques, l'argument de la direction d'ajustement et l'argument de l'uniformité théorique. Du côté anti-humien, l'objection voulant que la notion de direction d'ajustement soit circulairement caractérisée, l'objection des croyances irrationnelles et l'objection selon laquelle certains états mentaux présenteraient deux directions d'ajustement. Je conclus qu'aucune de ces objections ne devrait nous conduire à rejeter la théorie humienne, mais aussi que cette dernière trouve son principal appui du côté des considérations relatives à l'uniformité théorique et non du côté de l'argument de la direction d'ajustement comme le pensent plusieurs.

Mots clés : Philosophie, métaéthique, théorie humienne de la motivation, direction d'ajustement, uniformité théorique.

## **ABSTRACT**

Although remaining orthodoxy in philosophy of mind, the humean theory of motivation (according to which motivation always comes from a desire, never from a belief alone) has often been criticized in the last decades. Here I review the main arguments put forward both to support and reject this theory. On the humean side, I examine the argument from identical beliefs, the direction of fit argument and the argument from theoretical uniformity. On the anti-humean side, the objection that direction of fit characterizations are circular, the objection from irrational beliefs and the objection according to which there could be mental states with two directions of fit. I conclude that none of these objections should lead us to reject the humean theory, but also that this one finds its main support in theoretical uniformity considerations and not in the direction of fit argument as many think.

Key words : Philosophy, metaethics, humean theory of motivation, direction of fit, theoretical uniformity.

## TABLE DES MATIÈRES

INTRODUCTION.....	1
CHAPITRE 1 : Considérations préliminaires	
1.1 La théorie humienne de la motivation plus en détail .....	12
1.2 La phénoménologie de l'expérience morale.....	20
CHAPITRE 2 : Arguments en faveur de la théorie humienne de la motivation	
2.1 L'argument des croyances identiques	
- Argument .....	28
- « Overriding » et subsomption adéquate .....	30
2.2 L'argument de la direction d'ajustement	
- Notion et argument .....	33
- Interprétation dispositionaliste .....	37
- Interprétation normativiste .....	41
2.3 L'argument de l'uniformité théorique	
- Argument .....	47
- Motivation <i>sub specie boni</i> ? .....	53
Conclusion du chapitre .....	60
CHAPITRE 3 : Arguments à l'encontre de la théorie humienne de la motivation	
3.1 L'argument de la circularité et de la non-informativité .....	65
3.2 L'argument des croyances irrationnelles.....	70
3.3 L'argument des états mentaux de double direction d'ajustement	
- Argument.....	75
- Neurones miroirs et intentionalité .....	80
Conclusion du chapitre .....	85
CONCLUSION GÉNÉRALE .....	89
BIBLIOGRAPHIE .....	101

## **Introduction**

L'étude du phénomène de la motivation appartient de droit à la philosophie de l'esprit et à la psychologie philosophique. Le débat contemporain sur la motivation s'articule pourtant autour de questions pour la plupart liées à un tout autre domaine de la philosophie, à savoir celui de la philosophie morale. Les théoriciens questionnent notamment la nature de la relation entre motivation et jugement moral : Un jugement moral entretient-il une relation nécessaire avec la motivation, de sorte que si un agent juge sincèrement devoir (du point de vue moral) poser une action, alors celui-ci sera nécessairement motivé à poser l'action en question ? En supposant qu'un jugement moral soit l'expression d'un état mental de l'agent qui le formule, un jugement moral devrait-il alors être considéré comme l'expression d'un état mental ayant une capacité motivationnelle intrinsèque – un désir par exemple, ou une croyance d'un type particulier ? Exprime-t-il plutôt un état mental sans capacité motivationnelle intrinsèque mais dont l'association avec un état d'une autre espèce parvient à générer la motivation ? Et si une telle association est une condition nécessaire pour qu'un agent soit motivé, de quel type d'association s'agit-il ?

On n'a qu'à aborder la littérature pour constater que le débat sur le thème de la motivation est loin d'être clos. Non seulement les différents protagonistes ne parviennent pas aux mêmes réponses lorsqu'ils considèrent les différentes questions soulevées par ce débat ; ils paraissent parfois même ne pas s'entendre sur la nature réelle des enjeux qui lui sont liés<sup>1</sup>. Toutefois, de façon générale (mais nous verrons qu'il existe des exceptions), les théoriciens paraissent d'accord pour affirmer que le débat devrait être formulé selon les termes d'une conception plutôt standard de la psychologie humaine ; conception selon laquelle ce sont des états mentaux, principalement des croyances et des désirs, qui génèrent la motivation. Aussi la question de la nature des états mentaux motivationnel (sont-ils des croyances ? des désirs ? une combinaison des deux ? ou encore des états d'un autre type ?) semble-t-elle au centre du débat.

Précisons d'emblée que les questions liées au phénomène de la motivation ne se posent que dans le cas d'événements pouvant être décrits comme des manifestations de comportements *intentionnels*, c'est-à-dire orientés vers un but. Ce ne sont effectivement que les *actions*, au sens propre du terme, qui soulèvent le problème de la motivation. Les événements qui n'impliquent pas une forme quelconque d'intention de la part de l'agent (soit-elle inconsciente) ne comptent pas comme des actions et il n'y a pas de sens à tenter de les expliquer en termes de motivations de l'agent. Lorsqu'on demande la « raison » d'un tel événement, on demande en fait sa cause ; et le type de cause invoqué ne peut en aucun cas constituer un motif d'action pour un

---

1. On s'en convaincra en comparant par exemple Smith, 1994, chapitre 4, qui est considéré par plusieurs comme une des principales défenses de la THM, et Shafer-Landau, 2007, pp. 225-230, qui conçoit le débat sur la motivation de telle manière que Smith se retrouve du côté anti-humien.

agent<sup>2</sup>. Bref, on peut poser la question de ce qui a motivé un agent à repeindre le plafond de sa cuisine, mais on ne peut pas poser la question de ce qui l'a « motivé » à tomber de l'escabeau lorsque ce dernier s'est brisé ou à digérer le repas qu'il vient d'avalier.

Une fois ces détails précisés et pris en considération (rappelons qu'il s'agit d'interroger la nature des états mentaux motivationnels, que le débat s'articule principalement en termes de croyances et de désirs et que parler de motivation n'a de sens que pour les actions proprement dites), on peut alors aborder le problème de la motivation à l'aide d'une question plus précise : il s'agit de savoir si et dans quelle mesure croyances et désirs participent à la genèse l'action. C'est principalement autour de cette question que, sur la scène de la métaéthique contemporaine, se partagent les « humiens » (ou « néo-humiens ») et « anti-humiens ». Parmi les premiers, entre autres : Simon Blackburn, Lloyd Humberstone, Michael Smith et Nick Zangwill. Parmi les seconds : Jonathan Dancy, Margareth Little, Thomas Nagel, Thomas Scanlon, Fred Schueler et quelques autres encore.

Selon les partisans de la théorie humienne de la motivation (THM), la motivation résulte toujours d'une paire d'états mentaux composée d'une croyance et d'un désir. Au sein de cette paire, le désir se voit toutefois attribuer le rôle principal tandis que la

---

2. Certains lecteurs remarqueront sans doute que ce petit passage cache des difficultés considérables : celle de savoir si, en quoi et comment *causes* et *raisons* doivent être distinguées ; celle de l'explication de la « causalité mentale » ; celle de savoir s'il est possible de maintenir une distinction entre quelque chose comme une causalité agentielle et la causalité purement événementielle (sans agentivité d'aucune sorte) qui prévaut hors du domaine des comportements intentionnels. Et il y a d'autres difficultés encore. Évidemment, il n'y a pas de solutions qui fassent l'unanimité pour toutes celles-ci. Toutefois, dans une large mesure, les aspects du débat sur la motivation qui nous intéressent sont dans la littérature développés indépendamment de ces questions. De plus, rien dans l'état actuel du débat ne donne à penser que les réponses à ces questions devront nécessairement être autres que celles présumées par les différentes formes de théories de la motivation présentées ici.

croissance, elle, ne reçoit qu'un rôle accessoire. C'est en effet le désir qui, pour ainsi dire, exerce à lui seul toute la « force » motivationnelle. Quant à la croyance, elle est motivationnellement « inerte » : elle ne sert qu'à mettre en lumière les objets qui sont pris pour fins par le désir (en révélant leur nature, leurs caractéristiques, etc.) ou à déterminer quels sont les moyens qui doivent être mis en œuvre pour parvenir aux fins fixées par celui-ci. L'influence du philosophe David Hume est ici évidente et l'on comprend aisément que le nom de celui-ci, transformé en épithète, serve à désigner les théories contemporaines de la motivation qui accordent une telle prépondérance au désir dans l'explication de l'action. C'est que, pour Hume, « la raison ne peut être à elle seule un motif pour un acte volontaire »<sup>3</sup>. La croyance (une activité de la raison) et le désir (une des formes de la passion) jouent selon lui des rôles très différents. Alors qu'un désir est un état mental intrinsèquement motivationnel, une croyance est un état mental ayant pour *seules* fonctions de représenter les relations entre nos idées ou de représenter les relations entre les objets de l'expérience (ces fonctions sont en fait celles que Hume attribue à la raison, mais, la croyance étant une activité de la raison, ses propres fonctions ne peuvent être que celles-ci). La croyance joue un rôle essentiellement représentatif et son apport à la motivation n'est donc que de présenter à l'agent les moyens d'atteindre une fin préalablement posée par son désir. La croyance est « inerte » du point de vue motivationnel<sup>4</sup>.

---

3.Hume, 1739-1740, p. 522.

4.Il est à noter que les interprètes de Hume ne sont pas tous d'accord pour affirmer que la THM correspond à une position réellement endossée par le philosophe écossais. C'est que certains passages du *Traité de la Nature Humaine* semblent difficilement conciliables avec l'image, disons un peu simpliste, que l'on se fait généralement des idées de Hume sur la motivation (à ce sujet, voir notamment Little, 1997, p. 66). Ce n'est évidemment pas l'objet de la présente recherche que de déterminer si oui ou non Hume était humien au sens de la THM. Il est cependant certain que la THM correspond à une idée que nous nous faisons souvent de la philosophie de Hume et certain aussi que la THM est largement inspirée de certaines des affirmations de Hume, particulièrement au sujet du rôle de la raison.

Ainsi, les humiens soutiennent que, puisque la force motivationnelle est entièrement contenue dans le désir de l'agent, la présence d'un désir est *toujours nécessaire* et celle d'une croyance *jamais suffisante* à la genèse de la motivation. De plus, les partisans de la THM soutiennent que la croyance et le désir sont des « existences distinctes », c'est-à-dire qu'il est toujours possible pour un agent d'être dans l'un des deux états mentaux sans être dans l'autre (i.e. croire sans désirer et vice versa). Il serait ainsi impossible que l'acquisition d'une croyance par un agent puisse, d'elle seule, susciter l'apparition d'un désir (préalablement inexistant, et pas seulement non-manifesté) dans l'esprit de celui-ci. Ici encore, la raison d'adopter cette thèse (l'indépendance d'existence) est à chercher du côté des idées de Hume lui-même. Puisque pour Hume la fonction d'une croyance est uniquement représentative, l'introduction (ou la suppression) d'une croyance dans l'esprit d'un agent ne peut pas modifier son état motivationnel, à moins que la croyance en question ne soit liée à un désir existant préalablement. Une croyance qui serait ainsi capable d'influer sur la motivation, c'est-à-dire par le truchement de la production d'un désir « dérivé », ne serait tout simplement pas une croyance au sens où Hume l'aurait entendu.

Face à la THM, on retrouve les théories de la motivation dites *anti-humiennes*. On qualifie aussi parfois ces théories de *cognitivistes*<sup>5</sup> puisque, dans l'explication de l'action, elles attribuent le rôle principal à la croyance, qui est un état mental dit « cognitif » (on entend généralement par là que le contenu de cet état mental peut être vrai ou faux selon qu'il s'accorde ou non avec ce qu'il décrit. Le désir, quant à lui, est un état mental « conatif », c'est-à-dire qu'il a une fonction motivationnelle inhérente.)

---

5. Voir par exemple Little, 1997 ou Tenenbaum, 2005.

Les théories anti-humiennes prennent des formes diverses mais elles peuvent être regroupées en deux grandes familles : celles qui affirment qu'il existe des croyances pouvant générer la motivation sans le secours de désirs extérieurs et celles qui affirment qu'il existe des croyances telles que leur présence chez un agent entraîne nécessairement la présence d'un désir (de contenu approprié) chez ce même agent. Les théories de la première famille remettent en cause la thèse humienne selon laquelle le désir est toujours nécessaire et la croyance jamais suffisante pour engendrer la motivation. Quant à celles de la deuxième famille, elles s'en prennent plutôt à la thèse selon laquelle le désir et la croyance (formant la paire d'états mentaux nécessaire à la genèse de la motivation) sont toujours des existences distinctes.

La THM a longtemps constitué la théorie principale en ce qui concerne la nature des états mentaux motivationnels. La THM est certes une théorie importante en tant qu'elle est une des principales réponses offertes au problème de la motivation morale ; problème lui-même important et étroitement lié à un certain nombre d'autres problèmes pas moins importants : celui de l'éducation morale, celui de la responsabilité morale, celui de la possibilité d'une authentique action altruiste, etc. Mais l'importance de la THM sur la scène de la métaéthique contemporaine est aussi en grande partie due au fait que la position adoptée par rapport à la THM est déterminante quant à la façon de résoudre certaines des questions métaéthiques les plus fondamentales, notamment celle de la nature du jugement moral, celle du statut ontologique des entités morales et celle du lien entre jugement moral et motivation<sup>6</sup>.

---

6. Sur le fait que ces questions constituent en quelque sorte le cœur de la métaéthique et que la position adoptée face à la THM détermine (dans une certaine mesure) les réponses qu'on peut apporter à ces questions, voir Smith, 1994, pp. 126-129.

Supposons par exemple que l'on adopte la THM et que l'on adopte aussi la thèse de l'internalisme du jugement moral, selon laquelle il existe un lien nécessaire entre le fait de juger (sincèrement) qu'une chose est bonne (ou juste, ou désirable, etc.) et le fait d'être motivé à obtenir cette chose – si je crois que  $x$  est bon, alors je serai motivé à agir de manière à obtenir  $x$ . Il semble alors qu'on doive endosser la position non-cognitiviste en sémantique du jugement moral. En effet, si un jugement moral motive nécessairement celui qui le formule (internalisme du jugement moral) et que la motivation découle toujours d'un désir (THM), alors un jugement moral doit être l'expression d'un désir de l'agent – jamais l'expression d'une croyance seule – (non-cognitivism sémantique). De plus, si l'on croit que le réalisme moral « implique que les jugements moraux sont des croyances qui sont parfois vraies et qui, lorsque vraies, énoncent quelque chose de factuel », alors l'abandon du cognitivism sémantique devrait avoir pour conséquence le rejet de la position réaliste en ontologie morale<sup>7</sup>.

Supposons maintenant que l'on adopte la THM et une théorie cognitiviste du jugement moral, c'est-à-dire une théorie pour laquelle le jugement moral exprime une croyance prenant pour objets des faits moraux et pour laquelle le jugement moral peut être vrai ou faux. Il faudra alors que l'on adopte une conception *externaliste* du lien entre jugement moral et motivation : puisque seul un désir peut motiver (THM) et qu'un jugement moral est l'expression d'une croyance (cognitivism sémantique), être

---

7. Le passage cité est de Shafer-Landau (2003, p. 121). Je mets la phrase entière au conditionnel car on pourrait imaginer qu'il existe quelques préceptes moraux absolus (dans l'esprit de dieu par exemple), qui existent indépendamment du jugement de tout agent particulier et qui existent malgré le fait que le jugement d'un agent particulier n'exprime rien d'autre qu'une attitude conative qu'il a relativement à une chose donnée. Ce serait l'inverse d'une théorie de l'erreur (Mackie, 1977) : ce ne serait pas que nos jugements visent à rapporter des faits moraux qui n'existent pas, mais plutôt que les faits moraux existent sans que nos jugements ne visent à les rapporter.

motivé conformément à un jugement moral signifie nécessairement être motivé par un désir « externe » (le désir d'agir moralement par exemple).

À l'inverse, défendre simultanément le cognitivisme sémantique et l'internalisme du jugement moral, cela implique d'endosser une théorie anti-humienne de la motivation : si un jugement moral exprime une croyance (cognitivisme) et qu'un jugement motive nécessairement (internalisme), alors une croyance suffit à motiver (anti-humianisme).

Si la THM a longtemps constitué la théorie principale en ce qui concerne la nature des états mentaux motivationnels, le dogme que personne ne mettait en question, le nombre des opposants à celle-ci s'est considérablement accru au cours des dernières années. De sérieuses objections sont avancées à son encontre. Nous venons de voir que la THM est au cœur de problèmes théoriques majeurs, principalement en philosophie morale. Tenter de clarifier le débat entre humiens et anti-humiens, voire de prendre position au sein de celui-ci, est donc d'une importance considérable eu égard à la compréhension des problèmes en question. Les pages de ce mémoire seront dédiées à cette tâche. Mais, avant d'entrer dans le vif du sujet, voyons tout d'abord le plan que suivra ma recherche.

Je me propose ici de passer en revue quelques-uns des principaux arguments avancés en faveur et à l'encontre de la THM. Considérant le nombre de ces arguments et l'ampleur du traitement qu'il convient de donner à chacun pour lui rendre justice, il serait tout simplement irréaliste de viser l'exhaustivité dans le cadre d'une recherche d'une centaine de pages. J'ai donc choisi d'exposer les arguments qui sont le plus souvent discutés dans la littérature. Il me semble que c'est bien à travers la discussion

de ces arguments qu'il est possible, d'une part de saisir la nature du débat opposant les humiens aux anti-humiens et, d'autre part, d'évaluer les forces et les faiblesses des positions théoriques endossées par ceux-ci.

Mon mémoire s'organisera donc de la façon suivante. Le premier chapitre servira à présenter quelques considérations préliminaires à l'examen des arguments en faveur et à l'encontre de la THM. Ce chapitre se divisera en deux parties. La première nous permettra de préciser différentes notions et principes liés à la THM, de délimiter la portée explicative de celle-ci et aussi de montrer que la THM est au moins superficiellement plausible. Cette dernière tâche exigera notamment que l'on explique comment la THM peut s'appliquer à des cas qui, à première vue, semblent difficilement pouvoir recevoir une explication humienne. Quelques-uns de ces cas sont soulevés sur la base de considérations relevant de la phénoménologie de la motivation morale (Shafer-Landau, 2003). La deuxième partie de ce premier chapitre sera donc consacrée à un bref examen de la phénoménologie de la motivation morale. Suivra un chapitre destiné à l'exposition et à la discussion des arguments en faveur de la THM. Trois arguments seront alors examinés : l'argument des croyances identiques (Smith, 1994 ; Little, 1997), l'argument de la direction d'ajustement (Smith, 1994 ; Zangwill, 1998), et l'argument de l'uniformité théorique (Shafer-Landau, 2003). Quant au troisième et dernier chapitre, il tiendra lieu d'exposition des principales objections adressées à la THM et des réponses (s'il y en a) que peuvent faire les partisans de la THM à celles-ci. Les arguments examinés dans ce chapitre seront les suivants : l'argument de la circularité et de la non-informativité des tentatives faites pour distinguer les désirs des croyances à partir de la notion de direction d'ajustement

(Humberstone, 1992 ; Copp et Sobel, 2001), l'argument des croyances irrationnelles (Copp et Sobel, 2001) et, pour finir, l'argument des états mentaux de double direction d'ajustement (Pettit, 1987 ; Millikan, 1995 ; Little, 1997).

On le voit, une grande partie de ma recherche consistera à décrire et expliquer quelques-uns des moments stratégiques du débat entre humiens et anti-humiens. Une telle revue des arguments avancés en faveur et à l'encontre de la THM permettra de rassembler de façon claire et concise de nombreux éléments épars dans la littérature ; éléments dont la compréhension est d'une importance considérable puisqu'ils se retrouvent au cœur de nombreux débats philosophiques : le débat internalisme/externalisme, le débat cognitivisme/non-cognitivisme, le débat réalisme/anti-réalisme, le débat portant sur les rapports entre éthique et égoïsme, etc. Cependant, en plus de cet aspect descriptif, mon travail comportera un aspect critique. Je me propose en effet de présenter et d'examiner les différents arguments pro et anti-humiens dans le but de les évaluer. Je mettrai donc en lumière les forces et les faiblesses de ces arguments et tenterai de montrer que, en fin de compte, les objections adressées à la THM ne sont pas décisives et que cette dernière trouve un appui suffisant du côté des arguments avancés en sa faveur. Nous verrons en fait que, en l'état actuel du débat, des considérations d'économie théorique devraient nous conduire à préférer la THM à ses rivales ; la THM étant de ce point de vue plus parcimonieuse que les autres théories de la motivation.

En ce qui concerne la méthode de travail que je compte suivre, celle-ci consistera principalement en une revue de la littérature philosophique portant sur la théorie humienne de la motivation et en une analyse conceptuelle des notions en lien avec

celle-ci. Mais, étant donné le fait que les différentes théories de la motivation présupposent quelques fois des prises de position qui peuvent (ou non) trouver un appui du côté empirique – et bien que mon approche sera avant tout philosophique –, je me propose aussi de prendre en compte certains résultats pertinents de la psychologie empirique et des neurosciences (Arbib, 2007 ; Rizzolatti, 1988).

Tel qu'annoncé, nous débuterons maintenant par quelques considérations préliminaires à l'examen des arguments en faveur et à l'encontre de la théorie humienne de la motivation.

## **Chapitre 1**

### **Considérations préliminaires à l'examen des arguments en faveur et à l'encontre de la Théorie Humienne de la Motivation**

Comme je l'ai déjà dit, une section de ce premier chapitre servira à préciser davantage ce en quoi consiste la théorie humienne de la motivation. Ces précisions seront utiles à la compréhension des différents arguments en faveur et à l'encontre de la THM (chapitres 2 et 3). Cette section servira aussi à délimiter la portée explicative de la THM (en distinguant notamment le domaine des raisons motivantes de celui des raisons normatives) et à montrer que la THM peut s'appliquer à certains cas de motivation qui, à première vue du moins, semblent difficilement conciliables avec la théorie. Ainsi, bien qu'à proprement parler il n'y sera pas encore question des

arguments en faveur de l'adoption de la THM, cette première section devrait tout de même permettre de mettre en lumière quelques-unes des raisons qui font que la THM jouit d'une plausibilité au moins superficielle aux yeux de plusieurs théoriciens.

## 1.1 La THM plus en détails

Nous avons vu en introduction que la théorie humienne de la motivation consiste essentiellement dans l'affirmation selon laquelle la motivation d'un agent résulte toujours d'une paire d'états mentaux formée d'un désir et d'une croyance. Le désir porte sur une fin que l'action permettra d'atteindre, tandis que la croyance porte sur les moyens à mettre œuvre pour atteindre celle-ci. Le désir est le seul véritable moteur de l'action : sans lui, la croyance ne donne lieu à aucune motivation. Évidemment, bien qu'elle n'ait aucune force motivationnelle véritable, la croyance est nécessaire à la production de l'action puisque, sans elle, la force du désir n'aurait aucune « direction » dans laquelle s'exercer : le désir a besoin de la représentation des moyens d'atteindre la fin qu'il a fixée.

Ainsi, il existe un sens selon lequel la THM accorde une forme de primauté à la croyance dans le schéma explicatif de la genèse de la motivation : la croyance intervient en premier dans la mesure où elle doit *d'abord se manifester* dans l'esprit de l'agent et offrir au désir les moyens de sa satisfaction, pour qu'*ensuite* celui-ci

devienne pour ainsi dire effectif. Mais cette forme de primauté (dans la chronologie de la manifestation des états mentaux<sup>8</sup>) n'est évidemment pas celle sur laquelle la THM est reconnue pour avoir mis l'accent. Une des questions fondamentales qui opposent humiens et anti-humiens est celle de savoir à quel type d'état mental particulier (désir et/ou croyance et/ou état d'un autre type) doit être attribuée la capacité motivationnelle nécessaire à la production de l'action. Les humiens accordent la primauté explicative (à vrai dire, toute la valeur explicative) au désir, en tant que celui-ci exerce à lui seul toute la force motivationnelle à l'œuvre lors d'une occurrence du phénomène de la motivation. C'est donc l'importance relative du rôle du désir dans la genèse de l'action qui constitue la raison du primat accordé par les humiens au désir. La primauté temporelle de la croyance sur le désir n'indique pour eux rien de significatif quant à l'attribution correcte de la capacité motivationnelle à un type d'état mental particulier. (Nous verrons, au troisième chapitre, que les anti-humiens ont des opinions bien différentes relativement à la primauté temporelle de la manifestation de la croyance dans le schéma explicatif de la genèse de l'action.)

La prédominance du rôle accordé par les humiens au désir dans la genèse de l'action est le revers d'une médaille dont l'envers est, évidemment, une certaine conception de la croyance comme « inerte ». Cette conception est héritée de Hume lui-même, qui considérait que la raison a pour seules fonctions de juger, par démonstration, les relations abstraites qu'entretiennent nos idées entre elles et, par

---

<sup>8</sup>.Peut-être doutera-t-on de la primauté temporelle de la croyance, sous prétexte que, dans le cas de certains désirs « fondamentaux » (disons le désir de s'abreuver quand on est déshydraté), le désir semble devoir exister au moins sous forme de disposition et donc précéder la croyance (qu'il y a un verre d'eau ici par exemple) dans le schéma explicatif de la genèse de la motivation. Cela est juste. Mais il faut remarquer que je ne parle ici que de primauté *dans la manifestation* des états mentaux.

probabilité, les relations qu'entretiennent entre eux les objets de notre expérience. Ainsi, le rôle de la raison est essentiellement représentatif (représenter des relations d'idées ou des relations entre objets de l'expérience) et la raison « ne peut jamais à elle seule produire une action, ni engendrer une volition »<sup>9</sup>. Le débat contemporain est encore très influencé par cette conception de la raison – et donc de la croyance en tant qu'activité de celle-ci. À vrai dire, il semble que très peu d'auteurs confèrent aux croyances de fonctions autres que représentatives. Les humiens seraient en tous cas d'accord pour dire que, même s'il s'avérait que les croyances aient d'autres fonctions que celle de représenter les relations mentionnées, ces fonctions ne seraient pas telles que leur exercice puisse suffire à produire une motivation.

Il est sans doute vrai que si l'on considère des exemples paradigmatiques de croyances, il nous semble alors qu'il ne s'agit pas d'états mentaux susceptibles d'inciter par eux seuls un agent à l'action. Pensons à la croyance *que Madrid est en Espagne*, *que la gestation chez l'éléphant est de 21 mois* ou *que  $2a \times 3b = 6ab$* . Personne ne pense que des croyances de ce type peuvent d'elles-mêmes motiver un agent. Leur introduction dans l'esprit d'un agent n'engendrera une motivation que si elles sont accompagnées d'un désir du type approprié. Apprendre par exemple *que Madrid est en Espagne* ne peut générer une motivation, disons la motivation à aller à Madrid, que par association avec le désir (préalablement présent dans l'esprit de l'agent) d'aller en Espagne. Vraisemblablement, l'introduction de cette croyance parmi les états mentaux d'un agent n'induirait aucun changement au niveau de ses motivations si aucun de ses désirs ne dépendait (de près ou de loin) de Madrid ou de l'Espagne pour sa réalisation.

---

9.Hume, 1739-1740, p. 524.

Mais, s'il en va de la sorte pour les exemples paradigmatiques de croyances, le débat persiste quant à la question de savoir si tous les types de croyances sont ainsi inertes. Les anti-humiens remettent en question la thèse de l'inertie des croyances. Selon eux, les croyances *morales*, par exemple, pourraient être dotées de la capacité de motiver par elles-mêmes (ou de la capacité d'engendrer par elles-mêmes un état mental intrinsèquement motivationnel). Un autre type de croyance souvent considéré comme pourvu d'une telle capacité sont les croyances *évaluatives*, c'est-à-dire celles qui portent sur la valeur des choses (la croyance *que x est bon, que A est meilleur que B*, etc.). Nous aurons abondamment l'occasion de revenir à ces types de croyances.

Ainsi, d'une part la THM attribue un rôle prédominant au désir dans la genèse de la motivation et, d'autre part, elle s'appuie sur une conception de la croyance selon laquelle la fonction de la croyance est essentiellement représentative. Il ne faut pas oublier qu'il s'agit en fait de deux aspects d'un même mouvement conceptuel : c'est principalement parce qu'ils considèrent que la croyance est inerte que les partisans de la THM attribuent toute la force motrice au désir. Voilà pourquoi selon eux la présence d'un désir est toujours nécessaire et celle d'une croyance jamais suffisante pour qu'il y ait motivation. (Rappelons qu'il s'agit-là de deux des formules proposées en introduction pour caractériser la THM. Une caractérisation complète nécessiterait de rajouter que le désir et la croyance constituent des « existences distinctes ».)

Attardons-nous un instant sur une autre formulation proposée pour la THM. C'est peut-être en fait celle qui est la plus fréquemment utilisée lorsqu'il s'agit de décrire la THM d'une manière un peu plus formelle :

R à t constitue une raison motivante [*motivating reason*] de l'agent A de faire G si et seulement s'il existe quelque H tel que R à t consiste en un désir de A pour H et une croyance, reliée de manière appropriée, que s'il faisait G il obtiendrait H, où le désir et la croyance en question sont des « existences distinctes ». (Miller, 2003 : 271)

Il s'agit d'une proposition de Michael Smith<sup>10</sup>, défenseur notoire de la THM. Celle-ci met de l'avant la notion de « raison motivante » qui mérite un instant d'attention. Qu'est-ce en fait qu'une raison motivante ? La notion n'est pas souvent explicitée. Il ne s'agit pourtant peut-être pas d'une notion anodine. En voici une définition : « Une raison motivante est un état psychologique capable de jouer le rôle d'expliquer pourquoi un agent a agit ainsi qu'il a agit ; un état psychologique susceptible d'expliquer le comportement. »<sup>11</sup> Ce qu'il importe de remarquer c'est que, si cette définition est juste et qu'une raison motivante est « un état psychologique capable de jouer le rôle d'expliquer pourquoi un agent a agit ainsi qu'il a agit », alors une raison motivante n'est pas nécessairement quelque pensée présente à l'esprit de l'agent ayant la raison motivante en question. Dans la mesure où un état psychologique peut être inconscient, une raison motivante peut même échapper complètement à l'agent ayant cette raison. L'agent peut scruter attentivement ses désirs et, puisque le désir qui l'a poussé à l'action est inconscient, ne trouver nulle part celui qui pourrait expliquer l'action qu'il vient de poser. Il est incapable d'identifier la raison motivante qui l'a poussé à accomplir l'action qu'il a effectivement accomplie (ni même à vrai dire d'affirmer qu'il avait une telle raison) ; et donc incapable d'expliquer son action.

---

10. Smith, 1987, p. 36.

11. Miller, 2003, p. 271.

Les raisons motivantes doivent être distinguées d'un autre type de raisons : les raisons *normatives*<sup>12</sup>. La confusion de ces deux types de raisons pousse parfois à concevoir (à tort) certains cas de motivation (ou d'absence de motivation) comme des contre-exemples à la THM. En fait ces contre-exemples montreraient que la THM ne parvient pas à spécifier les conditions nécessaires et suffisantes dans lesquelles un agent peut être dit posséder une « raison d'agir ». Une raison normative est une raison qui rend une action intelligible en la *justifiant*, c'est-à-dire en montrant que celle-ci est obligatoire (ou bien permise) du point de vue d'un système normatif donné. Une raison motivante, pour sa part, rend une action intelligible en l'*expliquant*, c'est-à-dire en montrant que celle-ci résulte d'une cause dont la présence était (dans les circonstances de l'action) suffisante pour que l'action se produise. C'est ce deuxième type de raison seulement qui nous intéresse ici : nous nous questionnons sur la nature de l'état (ou des états) psychologique dont la présence dans l'esprit d'un agent est une cause suffisante de ce que celui-ci ait accompli l'action qu'il a effectivement accomplie.

La THM stipule que la présence d'un désir de contenu approprié est une condition nécessaire de la formation d'une raison *motivante* et non pas une condition nécessaire de l'existence d'une raison normative. Il existe sans doute un sens selon lequel lorsque j'écrase (accidentellement) les orteils d'une personne, *j'ai* alors *une raison* de retirer mon pied, et ce *en dépit de l'absence de tout désir de ma part* de retirer mon pied. Mais cela ne constitue pas un contre-exemple à la THM. Cela en constituerait un seulement si par « raison » on devait nécessairement entendre « raison motivante ».

---

12.Smith, 1994, p. 94 et sqq.

Mais c'est au sens de raison *normative* – qui rendrait mon action intelligible en déterminant son caractère obligatoire (ou permis) du point de vue d'un système normatif particulier – que j'ai une raison de retirer mon pied. De même, c'est seulement au sujet des raisons motivantes (et non des raisons normatives) que la THM affirme qu'elles ont pour condition suffisante la présence d'un désir et d'une croyance de contenu approprié. Ainsi, ce n'est pas aller à l'encontre de la THM que d'affirmer qu'il existe un sens selon lequel un agent *n'a pas de raison*, par exemple, de boire le poison placé devant lui, bien qu'il ait le désir de boire un certain breuvage et la croyance selon laquelle c'est ce breuvage (et non du poison) qui se trouve devant lui (donc même si l'agent a une raison motivante au sens de la THM) : cette raison que l'agent n'a pas est une raison normative (en l'occurrence prudentielle) et non une raison motivante.

La confusion entre raisons motivantes et raisons normatives est à la source de nombreux malentendus au sujet de la THM. Mais, comme nous venons de le voir, rien ne nous oblige à penser que les conditions de la genèse d'une raison motivante doivent être les mêmes que celles de l'existence d'une raison normative. Je dois maintenant toucher quelques mots au sujet de la conception du désir mise de l'avant par la THM.

Nous avons déjà vu que, pour rendre compte de certaines actions, la THM doit parfois recourir à des explications faisant intervenir des raisons motivantes partiellement constituées d'états psychologiques *auxquels l'agent n'a pas consciemment accès* – en particulier des désirs. Cette stratégie permet d'expliquer dans un cadre strictement humien des actions qui, à première vue au moins, paraissent ne pas pouvoir entrer dans un tel cadre. Mais voici une catégorie d'actions pour

laquelle la distinction entre désirs conscients et inconscients est impuissante à fournir une explication proprement humienne. Il s'agit des actions pour lesquelles la motivation ne peut être issue d'aucun désir *préexistant* pour un objet *particulier*.

La THM affirme que, pour qu'il y ait motivation, il faut qu'une croyance se joigne à un désir *préexistant* (et que les contenus de ceux-ci soient assortis évidemment). Mais pensons à la situation où un désir pour un objet donné semble naître au moment même où l'on apprend l'existence, les caractéristiques et le moyen d'obtenir de cet objet. Les croyances nouvellement acquises au sujet de cet objet ne peuvent pas alors se présenter à un désir *préexistant* de l'agent *pour cet objet singulier*. Imaginons par exemple une personne qui apprend l'existence d'un nouveau restaurant près de chez elle et qui ressent aussitôt le désir d'aller y manger. La raison pour laquelle, dans un tel cas, la motivation ne peut pas être issue d'un désir préexistant est que, évidemment, à l'instant précédant la genèse de la motivation, la personne ne connaissait rien du restaurant en question (pas même son existence). Elle ne pouvait donc pas avoir le désir préexistant (conscient ou inconscient) d'aller manger à *ce* restaurant et son désir semble alors avoir été généré par une croyance (en l'existence de ce restaurant, au sujet des caractéristiques de celui-ci, etc.). Mais on sait qu'il n'est pas possible pour un humien d'affirmer que la croyance nouvellement acquise a donné naissance au désir de la personne : cela irait à l'encontre du principe humien selon lequel le désir et la croyance formant la raison motivante sont des existences distinctes (i.e. qu'il est possible d'avoir l'un sans l'autre et vice versa). Le humien n'admet pas l'existence de quelque chose comme une croyance telle que son introduction dans l'esprit d'un agent suffirait à générer un désir chez celui-ci.

Cette difficulté peut toutefois être facilement levée si l'on fait porter le désir non pas sur l'objet particulier révélé par la croyance, mais plutôt sur le *genre* auquel appartient cet objet. Une personne ne peut certes pas désirer une chose particulière dont elle ignore l'existence, mais elle peut présenter un désir pour le type de choses exemplifié par cette chose particulière. En reprenant l'exemple précédent, on pourrait dire que le désir qui explique la motivation de cette personne à aller manger dans ce restaurant est en fait son désir (préexistant) plus général de manger dans des restaurants qui offrent un menu d'un certain type (auquel appartient le menu du nouveau restaurant).

La difficulté peut aussi être évitée pour certains désirs, à savoir ceux qui peuvent être considérés comme des dispositions. Si l'on reconnaît la présence de désirs chez les animaux, il semble alors plausible que certains de ces désirs (particulièrement ceux issus de leurs instincts) soient tels qu'ils ne se manifesteront jamais chez un spécimen qui n'a pas été exposé aux stimuli adéquats. Il en va probablement de même pour certains des désirs de l'être humain, en particulier ceux qui sont liés à ses besoins fondamentaux (manger, boire, dormir, avoir des rapports sexuels, etc.). Voilà qui peut, en partie au moins, expliquer que certains désirs semblent être engendrés par l'introduction d'une (ou plusieurs) nouvelle croyance dans l'esprit de l'agent chez qui surgissent ces désirs.

Avant d'aborder les arguments en faveur de la THM proprement dits, poursuivons notre examen préliminaire en nous attardant quelque peu à la phénoménologie de la motivation morale.

## 1.2 La phénoménologie de la motivation morale

Des considérations relatives à phénoménologie de la motivation morale sont parfois soulevées à l'encontre de la THM. Russ Shafer-Landau semblent même penser que le témoignage de la phénoménologie constitue un élément clé du débat sur la motivation ; élément qui, dans l'état actuel du débat, pourrait suffire à faire pencher la balance du côté anti-humien plutôt que humien<sup>13</sup>. Je reconnais qu'il peut être a priori intéressant d'aller voir ce que peut nous révéler la phénoménologie de la motivation au sujet de la nature de l'état mental (ou des états mentaux) présidant à la motivation, que ce soit la motivation morale ou non-morale. Je crois toutefois que la phénoménologie ne peut rien nous apprendre de décisif à ce sujet. Nous verrons bientôt pourquoi. Mais, si je crois que les considérations relatives à la phénoménologie ne sont en rien déterminantes dans le débat sur la motivation, je crois aussi que leur examen présente tout de même un intérêt en tant qu'il constitue l'occasion de montrer que la THM peut accommoder certains cas qui, à première vue, semblent peut-être lui échapper ; l'occasion donc de vérifier si la THM peut jouir d'une plausibilité au moins superficielle.

---

13. Shafer-Landau, 2003, p. 133.

Alors, qu'est donc supposée nous apprendre la phénoménologie de la motivation morale ? Lorsqu'on questionne les gens au sujet de ce qui les a motivé à accomplir une action donnée, il arrive bien sûr que ceux-ci citent la présence d'un désir parmi leur états mentaux. Mais il arrive fréquemment aussi que les gens ne réfèrent à aucun désir ni aucune autre attitude conative que ce soit. C'est souvent le cas avec les actions qui sont requises du point de vue moral. Les personnes affirment avoir agi moralement en l'absence de tout désir (ou autre attitude conative), simplement parce qu'elles « voyaient que c'était la bonne chose à faire » (i.e. elles croyaient cela). Les personnes affirment parfois même avoir agi moralement *en dépit* d'inclinations contraires : « Je voulais succomber, céder à mes désirs. [...] Je ne *voulais* pas faire x (me tenir devant cette balle, accepter le blâme, rester chaste), mais le devoir appelait. Je savais ce qui devait être fait, et je l'ai fait ; qu'importent mes désirs (ou passions, ou souhaits, ou inclinations). »<sup>14</sup>

Shafer-Landau appuie ces considérations en citant en exemples des cas où la motivation de l'agent semble découler de croyances *non*-morales. C'est ce qui arrive par exemple dans la situation simple où un agent considère des prémisses et en infère une conclusion. L'inférence semble alors commandée par les seules croyances de l'agent ; cela paraît étrange de supposer que l'agent devait alors avoir quelque chose comme le désir de faire l'inférence en question. Pareillement pour ce qui d'autres « actes mentaux » tels se souvenir, se concentrer, se soucier, etc. Il s'agit d'activités dont on voit mal comment elles pourraient constituer l'objet d'un désir. Elles semblent uniquement motivées par des croyances. Or, selon Shafer-Landau, il y aurait un sens

---

14. *Ibid.*, p. 123.

« à parler de nous-mêmes comme *faisant* de telles choses » ; ces actes mentaux sont pour lui de véritables actions. Shafer-Landau conclut : « Les désirs ne sont pas impliqués dans l'explication *la plus directe et la plus naturelle* de telles actions. »<sup>15</sup>

Un autre type de cas où la motivation d'un agent semble découler de ses croyances plutôt que de ses désirs, est celui où l'agent se méprend sur ses propres désirs et croit avoir un désir qu'il n'a pas en fait. L'exemple donné est celui d'un jeune homme qui commence à étudier le droit pour la raison (qu'il se donne à lui-même) qu'il croit qu'il désire une telle chose. Il constate bientôt que tout du droit l'ennuie et, après avoir travaillé avec joie comme charpentier pendant l'été, il décide d'abandonner définitivement le droit pour se consacrer à quelque chose de complètement différent. Après coup, le jeune homme se dit qu'il n'a jamais vraiment désiré étudier le droit, que c'est la croyance à l'effet qu'il désirait une telle chose qui l'a à l'époque motivé à entreprendre des études de droit. N'est-il pas alors (encore une fois) « plus naturel » et « plus direct » de soutenir que la croyance erronée de l'agent (à l'effet qu'il désirait étudier le droit) était véritablement l'état mental responsable de sa motivation ?

Shafer-Landau semble croire que le débat actuel sur la motivation est loin d'être résolu de manière définitive ; que les arguments en faveur et à l'encontre de chaque parti ne permettent pas de décider de l'issue du débat pour l'instant. Je suis assez d'accord pour dire que l'état actuel du débat ne permet pas d'en déterminer l'issue. Shafer-Landau croit toutefois que – puisque le poids des raisons est, pour ainsi dire, suspendu entre humanisme et anti-humanisme – la phénoménologie de la motivation morale révèle quelque chose de suffisamment important pour nous inciter à endosser

---

15. *Ibid.*, p. 124. Mes italiques.

une conception anti-humienne de la motivation : « nous avons quelques considérations, ayant à voir avec la phénoménologie de la motivation, etc., qui devraient nous pousser (légèrement, peut-être) vers l'anti-humianisme<sup>16</sup>. » Pour ma part, je pense que rien de ce que la phénoménologie révèle ne devrait nous inciter de la sorte à adopter une conception anti-humienne de la motivation.

Conie Rosati fait judicieusement remarquer que « les rapports introspectifs sont notoirement peu fiables et ne peuvent vraisemblablement pas résoudre une question si fondamentale de la psychologie morale. »<sup>17</sup> En ce qui concerne les cas où l'agent semble motivé par la croyance (erronée) qu'il a tel ou tel désir, rien n'indique que l'anti-humien dispose d'une explication « plus naturelle », « plus directe » que celle dont dispose le humien. Le humien peut en effet répliquer que c'est son explication qui est la plus simple et la plus conforme au sens commun. En reprenant l'exemple précédent, deux possibilités (au moins) s'offrent alors pour expliquer ce qui a motivé le jeune homme à étudier le droit et pourquoi sa motivation est disparue avec le temps. La première possibilité est simplement de dire que celui-ci désirait vraiment étudier le droit, mais qu'il avait une idée peu réaliste de ce en quoi consiste vraiment l'étude du droit, de sorte que son désir s'est éteint lorsque son objet s'est révélé sous ses traits réels. L'autre possibilité est d'affirmer qu'il ne désirait effectivement pas étudier le droit (comme une fin en soi), mais qu'il désirait autre chose reliée (par une relation de fin à moyens) à l'étude du droit ; et que c'est cela qui le motivait en fait. L'étudiant voulait par exemple, faire comme ses amis qui allaient étudier le droit, plaire à ses parents qui rêvaient de le voir exercer le métier d'avocat ou encore jouer

16. *Ibid.*, p. 133.

17. Rosati, 2006.

du prestige associé à la profession (mais il ne parle plus à ses amis, ses parents sont morts et le prestige associé à la profession est insignifiant en comparaison de la honte qu'il s'attire avec ses frasques).

Maintenant, en ce qui concerne les cas où l'agent semble motivé à établir une inférence, à se souvenir, ou à poser quelque autre acte mental de ce type ; je me contenterai de faire deux remarques. D'abord, il n'est pas clair qu'on ait en pareils cas toujours affaire à des actions au sens propre du terme. Cela est particulièrement évident dans le cas du souvenir et du souci : souvent, un souvenir nous vient à l'esprit sans que l'on cherche à se souvenir de quoi que ce soit (pas même inconsciemment). Se souvenir n'a alors rien d'intentionnel et ne doit pas compter pour une action. De même, c'est souvent de façon non intentionnelle que l'on se soucie de quelque chose. Il me semble que certaines inférences au moins sont établies de manière non-intentionnelle elles aussi : leur conclusion s'impose alors d'elle-même à l'esprit qui ne « cherchait » pas la réponse. Ensuite, s'il y a vraiment des cas où ces actes mentaux constituent des actions, je ne vois aucune invraisemblance dans la possibilité de désirer accomplir de telles actions. Il me paraît tout aussi possible de désirer se souvenir de quelque chose que de désirer manger un morceau de gâteau au chocolat. Les autres actes mentaux peuvent pareillement constituer l'objet d'un désir.

Pour résumer ce qui vient d'être dit au sujet de la phénoménologie de la motivation, une part des actes mentaux qui semblent initiés par des croyances doit être écartée du domaine des actions. Pour ceux de ces actes mentaux qui comptent comme des actions, rien n'interdit de penser qu'ils peuvent constituer de véritables objets de désirs. Quant aux actions qui paraissent motivées par une croyance erronée au sujet de

la possession d'un désir, elles peuvent recevoir une explication humienne tout aussi « naturelle » et « directe » que peut l'être l'explication anti-humienne : le humien peut dire que le désir de l'agent existait bel et bien, mais qu'il a disparu lorsque l'agent s'est mis à connaître davantage ce qui faisait l'objet de son désir ; ou alors dire que l'agent ne désirait pas en soi poser l'action qu'il a effectivement posée, mais plutôt qu'il désirait autre chose que cette action lui permettait d'atteindre (et que sa motivation a disparu en raison de la croyance nouvellement acquise à l'effet que cette chose n'était pas accessible par exemple). Mon avis est donc qu'il est juste d'affirmer que « ni la phénoménologie de la motivation morale ni les cas dans lesquels les individus se trompent au sujet de leurs désirs ne viennent appuyer la conception anti-humienne. »<sup>18</sup>

En terminant ce chapitre sur les considérations préliminaires à l'examen des arguments pro et anti-humiens, notons que d'autres précisions doivent être apportées relativement à la THM. Leur exposition devra pourtant être différée puisqu'elle exigerait que nous ayons déjà abordé quelques notions et thèses qui n'entreront en scène qu'au moment opportun dans le développement du présent mémoire. Certains détails relatifs à la THM seront donc présentés à mesure que les ressources conceptuelles nécessaires à leur appréciation auront été offertes au lecteur.

Il est maintenant temps de passer à la discussion des principaux arguments avancés en faveur de la THM.

---

18.Rosati, 2006.



## **Chapitre 2**

### **Arguments en Faveur de la Théorie Humienne de la Motivation**

Le présent chapitre aura pour objet l'examen de trois des principaux arguments avancés dans la littérature afin de soutenir la théorie humienne de la motivation. Ces arguments seront les suivants : l'argument des croyances identiques, l'argument de la direction d'ajustement et l'argument de l'uniformité théorique. Débutons avec l'argument des croyances identiques.

#### **2.1 L'Argument des Croyances Identiques**

##### **L'argument**

L'argument des croyances identiques est supposé offrir un appui à la THM en ce que, s'il réussissait, la plausibilité des théories anti-humiennes s'en trouverait sérieusement minée. En fait, l'argument vise à montrer que les théories anti-humiennes doivent être au moins en partie fausses puisque certaines de leurs implications sont fausses.

L'argument est très simple. Il consiste d'abord à faire remarquer que, puisque les théories anti-humiennes affirment que les croyances peuvent suffire à motiver un agent, ces théories impliquent que deux agents qui ont la même croyance seront identiquement motivés eu égard à l'objet de cette croyance. Or l'expérience (principalement dans le domaine de l'agir moral) nous apprend que parfois deux agents peuvent partager la même croyance sans être identiquement motivés. En fait, même un unique agent peut avoir la même croyance à deux moments distincts et être différemment motivé à chacun de ces moments. Il serait donc faux que les croyances peuvent à elles seules générer une motivation chez un agent.

De manière succincte :

- 1) Si les croyances suffisaient à motiver, alors deux personnes ayant la même croyance devraient être identiquement motivées.
- 2) Deux personnes partageant une même croyance peuvent être différemment motivées.
- 3) Donc les croyances ne suffisent pas à motiver.

On retrouve l'argument tel quel chez Shafer-Landau (qui ne le présente que pour en montrer l'inefficacité). On le retrouve aussi chez Smith, quoique sous une forme modifiée – individuelle plutôt qu'interindividuelle. Ce dernier considère les cas où un changement se produit au niveau de la motivation d'un agent sans qu'aucun changement n'ait lieu au niveau de ses croyances. De tels cas surviennent par exemple lorsqu'un agent souffre d'*akrasie* (ou « faiblesse de la volonté »), c'est-à-dire lorsqu'un agent n'est pas (ou pas suffisamment) motivé à poser une action donnée en dépit du fait qu'il juge (i.e. croit) que cette action est la meilleure des actions possibles. En diverses circonstances – dépression, fatigue, manque de concentration, apathie, désespoir, etc<sup>19</sup> –, il se peut donc qu'un agent perde sa motivation sans cesser pourtant de croire ce qu'il croyait lorsqu'il était motivé. C'est sans doute au sujet des actions requises du point de vue moral que l'on observe le plus souvent un tel phénomène : un agent croit que les principes moraux auxquels il adhère exigent qu'il pose une certaine action, mais n'est pourtant pas motivé conformément à cette croyance (en tous cas pas assez pour passer à l'action). Le même phénomène se constate aussi avec les actions jugées « supérieures » aux autres actions disponibles à partir de points de vue non-moraux (évaluatifs ou prudents par exemple).

Or, si les théories anti-humiennes étaient vraies et que la motivation d'un agent provenait de ses croyances, un tel phénomène serait tout simplement impossible, les croyances de l'agent restant les mêmes avant et après la perte de motivation. La différence qui existe entre un agent motivé et le même agent non-motivé ne tient donc pas à quelque croyance que ce soit (sinon bien sûr en tant que représentation des

---

19. Voir Stocker, 1979, p. 744.

moyens à mettre en œuvre pour atteindre les fins proposées par le désir).

Que l'on insiste sur la différence motivationnelle qui existe entre deux agents ou sur la différence qui existe chez un seul agent à deux moments différents, la stratégie argumentative reste la même : montrer que les théories anti-humiennes ont des implications fausses ; chose qui paraît peut-être de manière plus évidente à la lumière de ce que nous livre l'« expérience morale ordinaire »<sup>20</sup>.

### **Réplique Anti-Humienne : « Overriding » et Subsumption**

L'argument des croyances identiques offre en fait assez peu de soutien à la THM, voire pas du tout. Il existe en effet plusieurs manières de montrer qu'il ne constitue pas un réel problème pour la position cognitiviste. Shafer-Landau et Daniel Goldstick par exemple, attirent notre attention sur le fait que la motivation qui génère une action est, pour ainsi dire, la « somme vectorielle » des différentes motivations (concourantes et/ou concurrentes) présentes dans l'économie mentale de l'agent<sup>21</sup>. Une croyance ne doit pas être considérée seule, mais plutôt conjointement avec les autres croyances et désirs de l'agent. Il est donc possible qu'une croyance ait un réel pouvoir motivationnel mais que le pouvoir d'autres croyances ou d'autres désirs de l'agent soit plus fort et empêche que le pouvoir motivationnel de la croyance en question ne se transforme en action.

Prenons la croyance *que x est bon*. Selon la position cognitiviste, une telle croyance

---

20.Smith, 1994, p. 120.

21.Shafer-Landau, 2003, p. 129-130 ; Goldstick, 2001, pp. 159-160.

est du type approprié pour générer – sans le secours d'aucun désir – une motivation chez la personne qui l'entretient (disons la motivation à obtenir  $x$ ). Si la personne qui entretient cette croyance possède aussi un fort désir à l'égard de  $non-x$  (ou d'un aspect de  $non-x$ ), alors il se pourrait bien que la personne ne soit pas motivée à obtenir  $x$ . Sa motivation à obtenir  $x$ , quoique réelle et effective, est tout simplement « supplantée » (*overridden*) par la motivation issue du désir ayant pour objet  $non-x$ . De même, si la personne qui entretient la croyance *que  $x$  est bon* croit aussi *que  $y$  est meilleur que  $x$* , alors la motivation issue de cette autre croyance pourrait aussi supplanter la motivation à obtenir  $x$ .

Le fait qu'une croyance n'engendre pas la motivation lorsque présente avec d'autres états mentaux dans l'esprit de l'agent n'indique pas nécessairement que cette croyance n'est pas un état intrinsèquement motivationnel. En effet, du fait que le pouvoir motivationnel d'un désir puisse être supplanté par celui d'un autre désir, nous ne déduisons pas qu'un désir est autre chose qu'un état intrinsèquement motivationnel. Pourquoi devrait-il en aller autrement des croyances ? Et notons que même si deux personnes partageaient *l'ensemble de leurs croyances*, cela n'impliquerait pas que les deux doivent être identiquement motivées, puisque les désirs de ces deux personnes peuvent différer et que ceux-ci peuvent influencer sur la motivation issue de l'ensemble de croyances.

Margareth O. Little s'y prend d'une autre façon afin de montrer que l'argument des croyances identiques ne constitue pas un problème sérieux pour la position cognitiviste (anti-humienne)<sup>22</sup>. Little défend l'idée que la motivation qui résulte de la

---

22.Little, 1997.

croyance d'un individu (ou de plusieurs de ses croyances, voire de l'ensemble de celles-ci) peut différer selon que cette croyance est ou n'est pas subsumée sous un « cadre conceptuel plus large » (*broader conceptual framework*). Ce cadre conceptuel n'est pas lui-même une croyance ou un ensemble de croyances. Il s'agit plutôt d'une « manière de percevoir » la situation dans laquelle l'agent se trouve et cette manière de percevoir n'a rien à voir avec la saisie (ou l'absence de saisie) d'un ou plusieurs contenus propositionnels. Little établit une analogie entre ce processus de subsomption et le phénomène bien connu du « gestalt switch ». De même qu'un tableau pointilliste représentant le visage de Marilyn Monroe peut être perçu comme un ensemble de points (ne représentant rien) ou comme un visage, les caractéristiques d'une situation donnée peuvent être perçues comme étant, pour ainsi dire, « saillantes » ou comme étant dépourvues de cette qualité. Un agent peut conserver la croyance selon laquelle les points disposés sur le tableau représentent un visage tout en perdant temporairement la capacité de voir ce visage dans l'ensemble de points. Pareillement, un agent peut conserver la croyance *que x est bon* tout en cessant de percevoir cette caractéristique de *x* comme « moralement saillante ».

La différence pouvant exister au niveau de la motivation entre deux agents qui croient que *x* est bon est donc explicable par une différence cognitive entre ces deux agents. Bien que la différence au niveau de la motivation ne soit pas attribuée à une différence de croyances, cette objection à l'argument des croyances identiques reste néanmoins résolument anti-humienne, puisque la différence entre les deux agents est *entièrement* cognitive (aucun état ressemblant à un désir n'est ici en cause).

Il semble donc que l'argument des croyances identiques ne soit pas d'un grand secours pour établir la supériorité de la position humienne sur la position anti-humienne. Le fait que le pouvoir motivationnel d'une croyance donnée puisse être supplanté par celui d'autres croyances ou désirs de l'agent, ainsi que le fait qu'une croyance puisse ne pas être subsumée sous le cadre conceptuel approprié constituent des explications suffisantes de la différence pouvant exister entre les motivations de deux agents partageant une même croyance (ou plusieurs mêmes croyances, ou toutes les mêmes croyances). Pareillement pour ce qui est de la différence pouvant exister entre les motivations d'un unique agent à des moments distincts.

## **2.3 Argument de la direction d'ajustement**

### **La notion et l'argument**

La notion de *direction d'ajustement* (« direction of fit ») est presque incontournable dans la littérature portant sur la motivation. Désirs et croyances sont souvent dits se distinguer par leur direction d'ajustement. Mark Platts exprime cette idée de la manière suivante :

Les croyances visent le vrai et leur vérité consiste dans le fait qu'elles sont ajustées au monde (*their fitting the world*) ; être fausse signale un échec pour la croyance, et les croyances fausses doivent être abandonnées ; les croyances doivent être changées afin de s'ajuster au monde, non pas le contraire. Les désirs visent leur réalisation, et leur réalisation consiste dans le fait que le monde leur est ajusté (*the world fitting with them*) ; le fait que le contenu indicatif d'un désir n'est pas réalisé dans le monde ne signale pas en soi un échec du désir, et donc aucune raison en soi de rejeter ce désir ; le monde, grossièrement, doit être changé afin de s'ajuster à nos désirs, non pas le contraire<sup>23</sup>.

Donc, lorsque le contenu d'un état mental ne correspond pas aux faits du monde, si cet état est un désir, alors ce sont les faits du monde qui doivent changer ; dans le cas où l'état en question est plutôt une croyance, alors c'est l'état mental lui-même qui doit changer pour atteindre la concordance. On nomme parfois « monde-esprit » la direction d'ajustement associée au désir et aux autres états mentaux conatifs ; et, inversement, « esprit-monde » la direction d'ajustement associée à la croyance et aux autres états mentaux cognitifs.

Certains humiens ont tenté de recourir à la notion de direction d'ajustement dans le but d'élaborer un argument qui puisse montrer la nécessité de la présence d'un désir pour toute action intentionnelle. L'idée est d'identifier les propriétés dont la possession constitue le fait pour un état mental d'avoir l'une ou l'autre des directions d'ajustement. On montre ensuite qu'il est impossible à un seul état mental de posséder simultanément les deux directions d'ajustement : les propriétés correspondant aux directions d'ajustement sont telles qu'elles ne peuvent pas être instanciées

---

23.Platts, 1979, pp.256-257. Bien que celui-ci ait contribué à faire connaître une distinction principalement utilisée par les théoriciens humiens, Platts est lui-même défenseur d'une conception anti-humienne de la motivation. Notons aussi que c'est G.E.M. Anscombe qui fut la première à mettre de l'avant la notion de direction d'ajustement (voir Anscombe, 1963, p. 56).

simultanément par un même état mental (nous verrons bientôt pourquoi). Or, puisqu'une croyance possède essentiellement la direction d'ajustement esprit-monde (sans quoi on ne parlerait pas de croyance), une croyance ne saurait donc pas posséder aussi la direction d'ajustement monde-esprit. Il est tout simplement impossible pour une croyance de posséder la direction d'ajustement dont elle aurait besoin pour pouvoir motiver par elle seule. Pour qu'il y ait motivation, et donc action intentionnelle, il est nécessaire qu'un état mental ayant la direction d'ajustement appropriée soit présent ; un désir ou quelque autre état mental conatif.

Ce qu'il faut bien comprendre, c'est que les défenseurs de cet argument soutiennent qu'avoir la direction d'ajustement monde-esprit (la direction d'ajustement du désir) rend compte de l'action en tant que poursuite d'un but et que c'est cette dimension téléologique qui confère à l'action son intentionnalité<sup>24</sup>. Chez James Lenman et Michael Smith par exemple, l'argument de la direction d'ajustement prend la forme d'une explication du caractère intentionnel de l'action par le recours à la notion de direction d'ajustement du désir. L'argument qu'ils défendent peut être résumé ainsi :

- 1) Poser une action intentionnelle requiert de l'agent qu'il ait en tête quelque but ;
- 2) Avoir un but c'est être dans un état auquel le monde doit s'ajuster, et
- 3) Être dans un état auquel le monde doit s'ajuster c'est désirer.

(C) Donc, toute action intentionnelle nécessite la présence d'un désir<sup>25</sup>.

---

24. Sur la question de savoir si, en posant que toute action intentionnelle doit nécessairement présenter cette dimension téléologique, le humien n'est pas en train de biaiser le débat à son avantage, voir Shafer-Landau, 2003, p. 135.

25. Cette formulation est une interprétation faite par Schueler (2002, pp. 25-27) de Smith (1994, p. 116) et Lenman (1996, pp. 291-292).

G. F. Schueler a raison je crois d'affirmer qu'on pourrait vouloir établir la même conclusion en omettant tout bonnement les prémisses (2) et (3), c'est-à-dire en inférant celle-ci directement de la première prémisse. La tentation se fait sentir chez Lenman qui semble bien près d'affirmer que la direction d'ajustement monde-esprit est une caractéristique définissant (*defining characteristic*) l'action intentionnelle<sup>26</sup>. Mais les prémisses (2) et (3) jouent un rôle réel dans l'argument. Cela se reconnaît lorsqu'on comprend l'intention fondamentale de ceux qui tentent de distinguer désirs et croyances sur la base de leur direction d'ajustement.

Que veulent donc montrer les défenseurs de la THM en proposant un argument tel celui qu'on vient de voir ? On admettra sans doute que les actions intentionnelles sont dirigées vers quelque but déterminé par l'agent<sup>27</sup>. On admettra sans doute aussi que le fait d'avoir un but diffère considérablement du fait d'avoir une croyance dont le contenu est paradigmatiquement factuel (la croyance *que*  $5 + 7 = 12$ , *que l'Italie est en Europe*, etc.). Ainsi, si les défenseurs de l'argument de la direction d'ajustement font quelque chose « au-delà de répéter des truismes acceptés de tous dans une terminologie tortueuse – ainsi qu'ils le doivent, puisque leur conclusion n'est certainement pas acceptée de tous – ils doivent proposer une explication du fait qu'être orienté vers un but est différent de simplement avoir quelque sorte de croyance et s'avère nécessaire pour toute action ; une explication en termes d'états

---

26. Lenman, 1996, pp. 291-292, cité par Schueler, 2002, pp. 25-26. Comme le souligne Schueler, une telle chose ne serait que donner une définition stipulative de l'expression « direction d'ajustement » et alors rien de ce qui pourrait être établi à l'aide de la notion de direction d'ajustement ne pourrait l'être aussi à l'aide de la notion d'« avoir un but » (voir p. 27).

27. Tous les anti-humiens ne sont pas prêts à admettre une telle chose : voir la note 24.

psychologiques pouvant avoir une, et une seule, direction d'ajustement. »<sup>28</sup>

L'essentiel de la tâche qui attend le humien consiste alors, pour une large part, à caractériser de manière plus précise l'intuition derrière l'idée de direction d'ajustement. Si l'argument de la direction d'ajustement est souvent considéré comme l'argument le plus important en faveur de la THM, il est probablement aussi l'argument qui donne lieu aux considérations les plus subtiles sur la motivation. Cette subtilité se retrouve notamment dans la manière dont les théoriciens interprètent l'intuition derrière la notion de direction d'ajustement. Exprimée ainsi que nous l'avons vue, l'idée de direction d'ajustement n'est pas très claire. Il s'agit en fait d'une métaphore plus qu'autre chose. Qu'est-ce par exemple que de « devoir s'ajuster » pour une croyance ?

Pour donner un sens à la métaphore de la direction d'ajustement, les humiens ont élaboré des conceptions de deux types : d'une part, les conceptions qui donnent à l'idée de direction d'ajustement un sens purement descriptif (Smith, 1994 ; Lloyd Humberstone, 1992) et, d'autre part, les conceptions qui donnent à cette idée un sens normatif (Anscombe, 1957 ; Zangwill, 1998).

### **Le dispositionalisme de Smith**

Smith est le plus notoire défenseur de l'argument de la direction d'ajustement et son interprétation de la notion est la plus discutée dans la littérature sur la motivation. Cette interprétation s'appuie sur une conception particulière du désir, à savoir la

---

28.Schueler, 2002, p. 27.

conception dispositionaliste. Selon celle-ci, le désir consiste essentiellement en une disposition à accomplir quelque action ; action qui (sous une description donnée) est supposée permettre l'obtention de l'objet du désir. La conception dispositionaliste est choisie par Smith de préférence à la conception phénoménologique du désir. Cette dernière veut que les désirs soient comme les sensations, c'est-à-dire « simplement et essentiellement des états ayant un certain contenu phénoménologique »<sup>29</sup>. La conception phénoménologique du désir est rejetée par Smith pour avoir failli sur deux points : celui de la détermination des conditions correctes d'attribution d'un désir et celui de la détermination du contenu propositionnel d'un désir<sup>30</sup>. Pour ce qui nous concerne ici, il suffira de remarquer que la conception phénoménologique du désir, du moins telle que Smith se la représente, paraît incapable de fournir une interprétation crédible de la métaphore de la direction d'ajustement. Cette dernière affirme effectivement qu'un désir est un état qui tend à *persister* en dépit des indications à l'effet que son contenu correspond à un état de choses non-réalisé dans le monde. Or, dans le cas de certains désirs, il est difficile d'imaginer que ceux-ci persistent à travers le temps sous la forme de manifestations phénoménologiques. Smith donne l'exemple

---

29.Smith, 1994, p. 105.

30.Voici brièvement ce qu'il en est. La conception phénoménologique des désirs affirme que les désirs sont connus de la même manière que le sont les sensations, la douleur par exemple. Or, un agent ne peut pas se tromper relativement à la présence ou l'absence d'une douleur chez lui : si l'agent croit qu'il souffre, alors il souffre et, inversement, s'il croit qu'il ne souffre pas, alors il ne souffre pas. Nos intuitions les plus banales nous indiquent qu'il est possible pour un agent de se tromper au sujet des désirs qu'il possède : il existe des désirs inconscients et l'on croit parfois désirer une chose sans que cela ne soit réellement cas. La conception est donc fautive sur le plan des conditions correctes d'attribution d'un désir. De plus, les sensations sont typiquement des états sans contenu propositionnel. Une explication en termes de désirs de l'agent ne peut toutefois pas faire l'économie d'un tel contenu puisque ce qui fait que mon désir peut servir à expliquer mon action d'ouvrir la fenêtre, c'est le fait qu'il s'agit du désir *de me rafraîchir* (plutôt que le désir *d'éteindre la lumière*). La conception phénoménologique ne serait donc pas en mesure de déterminer le contenu propositionnel des états mentaux qu'elles est censée expliquer.

du désir d'un père pour le bien-être de son enfant. Un tel désir ne présente aucune manifestation phénoménologique en dehors d'épisodes précis : lorsque le père réfléchit plus intensément à l'avenir de son enfant, lorsque l'enfant est malade, etc. Il semble pourtant que ce désir existe de façon permanente, c'est-à-dire même en dehors de ces épisodes où il y a manifestations phénoménologiques.

Selon Smith donc, la conception phénoménologique du désir fait face à de graves difficultés et ne permet pas d'offrir une interprétation plausible de l'idée de direction d'ajustement<sup>31</sup>. Quelle est alors l'interprétation offerte par la conception dispositionaliste du désir ? L'idée que la croyance est telle qu'elle doit s'ajuster aux faits du monde tandis que le désir est tel que ce sont les faits du monde qui doivent d'ajuster à lui, devient avec le dispositionalisme l'idée que la croyance est un état qui tend à disparaître à la perception de la non-réalisation de son contenu tandis que le désir est un état qui tend à persister en dépit de la perception de la non-réalisation de son contenu et qui dispose l'agent à faire en sorte que son contenu soit réalisé. Cette interprétation de la notion de direction d'ajustement est essentiellement fonctionnaliste, c'est-à-dire qu'elle définit un état mental avant tout par ses relations avec d'autres états mentaux. Ainsi, la différence entre une croyance *que p* et un désir *que p* équivaut à « une différence au niveau de la dépendance contrefactuelle d'une croyance *que p* et d'un désir *que p* à l'égard d'une perception dont le contenu est non-*p* »<sup>32</sup> : alors que la croyance *que p* tend à disparaître à l'occurrence d'une perception de contenu non-*p*, le désir *que p* tend à persister à l'occurrence de cette même perception – et

---

31.Comme nous le verrons au prochain chapitre, la conception dispositionaliste n'est peut-être pas en meilleure posture.

32.Smith, 1994, p. 115.

dispose l'agent à agir de manière à obtenir p (nous verrons bientôt pourquoi cette dernière caractéristique du désir ne doit pas être omise).

La raison pour laquelle une croyance ne saurait motiver est qu'il est impossible à un même état mental de présenter simultanément les deux directions d'ajustement. Cela nécessiterait en effet que cet état manifeste des dispositions incompatibles : tendre à persister à l'occasion d'une perception et tendre à disparaître à l'occasion de cette même perception. La direction d'ajustement de l'état mental dans lequel se trouve un agent motivé est nécessairement la direction monde-esprit (le monde « doit s'ajuster » au contenu de cet état). Pour servir les fins de la motivation, un état mental doit en effet persister lorsque survient une perception à l'effet que l'état de choses désiré n'est pas réalisé. Or la croyance possède la direction esprit-monde de façon *essentielle* – sans quoi on ne parlerait pas de croyance. Une croyance ayant la capacité de motiver devrait donc avoir les deux directions d'ajustement et donc manifester simultanément une disposition à disparaître à l'occurrence d'une perception de contenu contraire à son propre contenu et une disposition à persister à l'occurrence de cette même perception. Cela est bien sûr impossible. Seul un désir (ou un autre état conatif) peut posséder la direction d'ajustement appropriée pour motiver.

On peut se demander si la présentation offerte par Smith de l'argument de la direction d'ajustement ne s'éloigne pas quelque peu de l'intuition à la source de la notion. En effet, l'idée était avant tout celle d'un certain type de relation entre des attitudes propositionnelles *et les faits du monde*. Or Smith pose que les relations recouvertes par cette notion sont plutôt des relations (causales) entre des attitudes propositionnelles *et d'autres attitudes* (par nécessairement propositionnelles cette

fois). Les faits du monde ne figurent pas directement dans l'argument. Mais, comme le disent David Copp et Mark Sobel, il va de soi qu'il ne peut pas y avoir de tendance pour les croyances à disparaître à l'occurrence de faits contraires « à moins que ces faits ne soient “pris à bord” cognitivement »<sup>33</sup>. Ce déplacement opéré par Smith (des faits du monde aux perceptions prenant sur ceux-ci pour objets) s'avère ainsi nécessaire. Donc l'idée de mettre en relation attitudes propositionnelles et faits du monde est conservée par l'interprétation de Smith ; seulement cette mise en relation est médiatisée par d'autres attitudes.

Précisons que l'interprétation donnée par Smith à l'argument de la direction d'ajustement vise à caractériser des relations causales entre, d'une part, les croyances et les désirs et, d'autre part, les faits (perçus) du monde. L'interprétation offerte de Smith se veut donc entièrement *descriptive* (ou « non-normative »), au sens où elle vise seulement à expliquer comment *sont* les choses, et non comment celles-ci *devraient être*. La deuxième interprétation que nous allons voir vise plutôt ce dernier objectif.

### **Le fonctionnalisme normatif de Zangwill**

Nick Zangwill défend une interprétation de la notion de direction d'ajustement qui, contrairement à celle de Smith, respecte une certaine intuition qui était à l'origine

---

33.Copp et Sobel, 2001, p. 46.

suscitée par la notion. La notion de direction d'ajustement, rappelons-le, fût d'abord mise de l'avant dans le but de rendre compte d'une relation *normative* entre les attitudes propositionnelles et les faits du monde. Ainsi G. E. M. Anscombe affirmait que la vérité de la croyance et la réalisation d'une intention se distinguent par la localisation de l'« erreur » (*mistake*) responsable de la non-concordance entre le contenu de l'attitude et les faits du monde. (Évidemment, le humien lit ici « désir » au lieu d'« intention ».) Anscombe imaginait un homme s'aidant d'une liste d'articles pour faire ses courses et un détective prenant en note quels articles sont achetés par l'homme en question. La liste du premier représente des intentions, tandis que la liste du second représente des croyances. L'absence de concordance entre le contenu de chaque liste et le contenu du panier de l'homme sera, *en tant qu'erreur*, localisée différemment : « si la liste et les choses que l'homme achète effectivement ne concordent pas, et si cela et cela seul constitue une *erreur*, alors l'erreur n'est pas dans la liste mais dans la performance de l'homme [...] tandis que si les notes du détective et ce que l'homme achète en fait ne concordent pas, alors l'erreur est dans les notes. »<sup>34</sup>

Zangwill soutient que l'interprétation de Smith (et, plus généralement, toutes les explications fonctionnalistes qui s'inspirent de la *folk psychology*<sup>35</sup>) ne livre pas l'essence même des attitudes de croyance et de désir ni donc l'essence de ce qui distingue celles-ci. Les dispositions sur lesquelles Smith attire l'attention ne sont, selon Zangwill, pas essentielles à ces attitudes : « L'essence d'un état mental explique ses pouvoirs causaux ; *elle n'est pas constituée par ceux-ci.* »<sup>36</sup> En fait, les théories du

34. Anscombe, 1957, p. 56.

35. Voir Zangwill, 1998, p. 182. Précisons que les auteurs ouvertement critiqués par Zangwill sont non seulement Smith, mais aussi Humberstone et Hume lui-même.

36. *Ibid.*, p. 179. Mes italiques.

genre de celle de Smith renversent le rapport explicatif qui devrait normalement s'établir entre la nature d'une attitude et les dispositions manifestées par celle-ci. C'est que les dispositions associées à chaque attitude découlent d'une nature qui n'est pas caractérisable selon les seuls termes d'une conception dispositionaliste. Zangwill fait en effet valoir deux arguments, celui de l'irrationalité systématique et celui des chaînes causales déviantes, afin de montrer qu'aucune forme de fonctionnalisme associée de trop près à la *folk psychology* (comme c'est le cas pour le dispositionalisme de Smith) ne pourrait saisir entièrement la nature de la croyance et du désir (ou de leur distinction). Tout fonctionnalisme de cette forme serait en effet impuissant à rendre compte du fait que certaines attitudes propositionnelles entretiennent des relations *rationnelles* avec d'autres attitudes propositionnelles et donc à rendre compte du caractère *normatif* des attitudes en question.

Le premier argument de Zangwill, celui de l'irrationalité systématique, rejoint une objection de Saul Kripke<sup>37</sup> à l'égard des conceptions dispositionalistes de la signification linguistique. Kripke a défendu l'idée selon laquelle la signification des termes et des énoncés d'un langage présente un élément de normativité inhérent : si je signifie vraiment l'addition par « + », alors je *dois* répondre « 125 » à la question «  $68 + 57 = ?$  » ; « 125 » est la seule réponse *correcte* à cette question. Le fait que je sois *disposé* à répondre cela ne peut pas en soi déterminer la signification du symbole « + » : je peux en effet être disposé à commettre certaines erreurs et ce parfois de façon systématique. En pareil cas (de disposition à se tromper systématiquement), si vraiment la fonction d'addition se définit par les équations que je suis disposé à

---

37.Kripke, 1982.

accepter, alors il est clair que je ne signifie pas l'addition lorsque j'emploie le terme « plus », le symbole « + », etc.<sup>38</sup> Les normes d'utilisation liées au sens d'un terme ou d'un énoncé semblent tenir malgré l'absence de dispositions adéquates. Le dispositionalisme (sémantique) est une théorie *descriptive* de la relation entre un terme (ou symbole, ou énoncé...) et son usage effectif. Il laisse par-là même échapper l'aspect *normatif* de la signification linguistique – aspect révélé par le fait que les dispositions mènent parfois à ce que l'on peut à juste titre appeler des « erreurs » (éventuellement systématiques). Pareillement pour ce qui est des conceptions dispositionalistes de l'essence des états mentaux : celles-ci laissent de côté le fait que certains états restent soumis à des normes bien que les dispositions manifestées par ces états puissent dévier systématiquement des normes en question.

Quant au deuxième argument de Zangwill, celui des chaînes causales déviantes, disons seulement à son sujet qu'il vise à écarter des formes plus complexes de fonctionnalisme ; plus précisément les formes *holistiques*, qui considèrent que l'irrationalité ne peut survenir que localement, « contre le fond d'autres relations causales systématiques qui sont maintenues ailleurs dans le réseau de dispositions »<sup>39</sup>.

En tant que théories de l'essence du mental, les théories fonctionnalistes sont donc fausses... « à moins que l'on ajoute les termes “rationalité” ou “justifié” dans les équations fonctionnalistes. Tout est alors dans l'ordre. »<sup>40</sup> Des théories fonctionnalistes du type approprié, c'est-à-dire celles qui reconnaissent l'aspect intrinsèquement

---

38.Kripke, 1982, p. 37.

39.Zangwill, 2003, p. 185. Si je n'explique pas véritablement l'argument en question et me contente d'en invoquer le but, c'est que son examen nous entraînerait trop loin du sujet qui est le nôtre ici. De toute manière, rien de ce qui suit ne dépendra de cet argument.

40.*Ibid.*, p. 181.

normatif des états mentaux, peuvent donc représenter adéquatement la nature des états mentaux. Zangwill qualifie d'ailleurs sa propre position de « fonctionnalisme normatif ». L'idée principale de ce fonctionnalisme est que croyances et désirs ont des liaisons normatives particulières. Une norme qui pourrait s'appliquer de manière plausible à certaines croyances serait par exemple que celles-ci sont rationalisées par les expériences perceptuelles d'un certain type. De même, une norme envisageable pour les désirs serait que ceux-ci rationalisent l'intention d'accomplir une action donnée ou rationalisent cette action elle-même.

La normativité des états mentaux est celle de la rationalité en général. C'est pourquoi les dispositions attribuées par Smith aux croyances et aux désirs (tendance à disparaître à l'occurrence d'une perception de contenu contraire dans le cas des croyances ; tendance à persister à l'occurrence d'une perception de contenu contraire et disposition à réaliser l'état de choses visé dans le cas des désirs) ne se manifestent pas de manière systématique. Elles se manifestent *dans la mesure où l'agent qui se trouve dans ces états mentaux est rationnel*. (Nous verrons au prochain chapitre qu'il s'agit là d'un sérieux problème pour l'interprétation proposée par Smith.)

Comment donc le fonctionnalisme normatif caractérise-t-il la distinction censée exister au niveau des directions d'ajustement de la croyance et du désir ? Zangwill divise les relations normatives entretenues par ces attitudes en « verticales » et « horizontales ». Les premières s'établissent entre une attitude et les faits du monde<sup>41</sup>. Les deuxièmes s'établissent plutôt entre deux attitudes. De l'avis de Zangwill, aucune distinction n'existe entre croyances et désir sur le plan des normes verticales qui leur

41. La vérité (dans le cas de la croyance) et la satisfaction (dans le cas du désir) pourraient constituer de telles normes.

sont liées<sup>42</sup>. La distinction recherchée se situe entièrement dans la manière dont chaque attitude rationalise une ou plusieurs autres attitudes :

L'asymétrie que nous considérons, ainsi, est que le désir que p, par lui-même, rationalise l'intention de faire en sorte que p, mais la croyance que p ne rationalise pas cette intention. Et l'expérience perceptuelle que p, par elle-même, rationalise la croyance que p, mais ne rationalise pas le désir que p. (Zangwill, 2003 : 195)

Pour les désirs et les croyances, se distinguer par leurs directions d'ajustement cela veut dire avoir différentes relations de rationalisation avec, d'une part, les intentions (qui sont rationalisées par les désirs mais pas par les croyances) et, d'autre part, les expériences perceptuelles (qui rationalisent les croyances mais pas les désirs).

Pour ce qui nous intéresse ici, préciser davantage les différentes interprétations qui peuvent être données de la notion de direction d'ajustement ne serait que de peu d'utilité. En effet, ce que nous avons vu ici à leur sujet suffira amplement à montrer qu'elles sont sujettes aux deux mêmes objections.

De tout ce qui vient d'être dit au sujet de l'argument de la direction d'ajustement, nous retiendrons donc que désirs et croyances sont censés se distinguer par leur « direction d'ajustement », c'est-à-dire que, tandis que la croyance présente un contenu qui doit s'ajuster aux faits du monde, le désir présente un contenu auquel les faits du monde doivent s'ajuster. Smith propose une interprétation descriptive de la notion de direction d'ajustement en adoptant une conception dispositionaliste des états mentaux.

---

42.Zangwill, p. 194.

Que les faits du monde « doivent s'ajuster » au contenu du désir, cela veut dire que le désir est un état qui persiste à l'occurrence d'une perception de contenu contraire et qui dispose l'agent à réaliser le contenu de son état mental. Qu'une croyance « doive s'ajuster », cela veut dire qu'elle tend à disparaître à l'occurrence d'une perception de contenu contraire. Un désir est donc nécessaire pour motiver un agent à l'action puisqu'une croyance ayant le pouvoir de motiver est un état qui, en principe même, ne peut pas exister : un tel état devrait manifester simultanément la disposition à persister et la disposition à disparaître à l'occurrence d'une perception de contenu contraire. Zangwill propose quant à lui de donner un sens normatif à la notion de direction d'ajustement : le contenu d'un état *doit* vraiment s'ajuster aux faits du monde (ou inversement, les faits du monde *doivent* s'ajuster au contenu d'un état). Cette normativité est celle de la rationalité elle-même. Ce qui distingue désirs et croyances c'est en fait qu'une perception peut rationaliser une croyance sans pouvoir rationaliser un désir et qu'un désir peut rationaliser une intention (ou une action) alors qu'une croyance ne le peut pas. La raison pour laquelle seul un désir peut motiver reste la même que pour Smith : un état cognitif tel une croyance devrait manifester simultanément des dispositions incompatibles s'il avait la capacité de motiver. Toutefois, contrairement à ce qui semble être le cas pour Smith, Zangwill considère que les dispositions liées aux états mentaux ne contribuent pas à définir ceux-ci. Ces dispositions *découlent* plutôt de la nature – intrinsèquement normative – des états mentaux en question.

Examinons maintenant ce qui, dans l'état actuel du débat sur la motivation, constitue selon moi le meilleur argument que l'on puisse avancer en faveur de

l'adoption de la position humienne.

## 2.3 L'argument de l'uniformité théorique

### L'argument

À ma connaissance, l'argument de l'uniformité théorique n'a jamais été employé par les humiens eux-mêmes pour appuyer leur théorie de la motivation favorite. Ce sont surtout les théoriciens anti-humiens qui, dans la littérature portant sur la motivation, ont fait connaître cet argument – pour le réfuter évidemment. On le retrouve notamment chez Shafer-Landau (2003 : 131-133) et sous forme de suggestion chez Little (1997 : 66) et chez Smith (1988 : 593).

En fait, l'argument vise à montrer qu'il existe une bonne raison de refuser une explication anti-humienne dans le cas spécifique de la motivation liée aux actions morales. C'est que l'argument *part* de la prémisse selon laquelle la THM est la théorie qui prévaut dans les cas d'actions non-morales : la motivation liée à ces actions serait mieux expliquée par la THM que par les théories rivales. Mais bien que l'argument parte ainsi d'une prémisse orientée d'emblée vers la position humienne et, pour cela, donne peut-être l'impression de devoir être écarté, nous verrons que son examen

permettra de prendre en compte certaines considérations qui, à mon avis, incitent fortement à penser que la THM doit effectivement prévaloir dans le cas des actions non-morales.

Donc, une fois posé que la THM prévaut dans le cas des actions non-morales, il s'agit de faire remarquer qu'expliquer la motivation liée aux actions morales en termes anti-humiens équivaldrait à constituer une théorie *non-uniforme* de la motivation. Il y aurait en effet d'une part les actions non-morales issues d'une motivation qui doit être expliquée à l'aide d'une théorie humienne et, d'autre part, les actions morales issues d'une motivation qui doit être expliquée à l'aide d'une théorie anti-humienne. L'argument consiste alors à affirmer qu'une théorie uniforme de la motivation est préférable à une théorie expliquant la motivation différemment selon que celle-ci est liée à une action morale ou non-morale. C'est que, évidemment, une théorie uniforme est plus parcimonieuse qu'une théorie non-uniforme et, *ceteris paribus*, la parcimonie est quelque chose de souhaitable du point de vue théorique.

Que les théories « hybrides » soient moins parcimonieuses qu'une théorie humienne (générale) de la motivation s'explique bien sûr par le nombre des explications requises pour rendre compte de la motivation dans les différents cas (moraux et non-moraux). Mais cela tient peut-être avant tout au fait que ces théories hybrides requièrent (au minimum) d'attribuer aux croyances une propriété dont on pourrait se passer en adoptant une théorie générale de la motivation.

Si je précise : « au minimum », c'est que le coût de l'adoption d'une théorie hybride est différent selon que l'on se place du point de vue humien ou anti-humien. Dans la perspective humienne, une croyance est essentiellement et uniquement un état

représentatif. Il est alors hors de question pour un humien d'attribuer la propriété de motiver à une croyance et l'état invoqué par l'anti-humien sera considéré comme étant d'un type autre. Du point de vue humien donc, le coût de l'adoption de la position anti-humienne est l'introduction d'un nouveau *type* d'état mental dans notre répertoire (et non pas la seule attribution d'une nouvelle *propriété* à un état d'un type déjà existant). Quoique certains théoriciens anti-humiens pensent que leur conception de la motivation implique effectivement de poser l'existence d'un état mental d'un type nouveau, il semble qu'il ne leur soit pas nécessaire de voir les choses ainsi. Il est envisageable pour eux d'attribuer à la croyance la propriété de motiver par elle-même ou celle de générer par elle-même un état mental intrinsèquement motivationnel. L'état mental posé par l'anti-humien est alors toujours considéré comme une croyance ; celle-ci est seulement dotée d'une propriété supplémentaire (celle de motiver ou celle de générer un état intrinsèquement motivationnel). Les deux possibilités sont offertes au théoricien anti-humien et sont bien illustrées dans la littérature par l'hypothèse du « croisir » et celle de la « desire-entailing-belief ». (Nous discuterons ces deux hypothèses, en particulier celle du croisir, au chapitre consacré aux objections à la THM, à la section portant sur les états mentaux de double direction d'ajustement).

Mais revenons à l'argument de l'uniformité lui-même. On pourrait je crois le résumer de la manière suivante :

1) La théorie humienne est celle qui prévaut dans l'explication de la motivation liée aux actions non-morales.

2) Introduire la croyance comme source de la motivation dans le cas des actions morales équivaut à introduire un nouveau type d'état (ou une nouvelle propriété) dans notre répertoire des états mentaux et à élaborer une conception non-uniforme de la motivation,

3) Par souci d'économie théorique (la parcimonie en ce domaine étant, *ceteris paribus*, quelque chose de souhaitable), il vaudrait donc mieux s'en tenir à la THM pour expliquer la motivation dans le cas des actions morales.

Qu'une théorie soit plus parcimonieuse qu'une autre est sans doute quelque chose de souhaitable, mais cela ne constitue pas une raison *décisive* de favoriser les théories uniformes dans le débat sur la motivation. C'est seulement lorsque toutes autres choses sont égales qu'une théorie parcimonieuse doit être choisie de préférence à une théorie qui l'est moins. Il faut donc tenir compte des autres raisons qui existent en faveur et à l'encontre de chaque type de théories (uniformes et non-uniformes) avant de déterminer si l'argument de l'uniformité théorique suffit à favoriser la THM dans le débat sur la motivation. C'est en partie la tâche à laquelle nous nous adonnons depuis le début du présent chapitre et à laquelle nous nous adonnerons au prochain chapitre avec notre analyse des principaux arguments en faveur et à l'encontre de la THM. Tant que nous n'avons pas considéré ceux-ci, nous ne pouvons pas savoir quelle sera ultimement l'importance de l'argument de l'uniformité théorique.

Mais laissons ces arguments précis de côté. J'ai dit en introduction qu'il existait des arguments en faveur et à l'encontre de la THM en dehors de ceux qui sont considérés ici. Ces arguments méritent sans doute une certaine attention mais, si le manque

d'espace m'a décidé à les laisser de côté, c'est parce qu'ils sont selon moi d'une moins grande efficacité que les arguments discutés ici. Je crois que l'argument de l'uniformité théorique ne souffrirait ni ne bénéficierait de la mise à contribution de ceux-ci. Mais peut-être de nouveaux arguments pourraient-ils surgir et changer la donne. Si donc, dans la situation présente, le humien peut grâce à l'argument de l'uniformité théorique prendre l'avantage dans le débat sur la motivation, l'avantage en question ne devra pas être considéré comme définitivement acquis.

Ainsi, il se pourrait qu'il existe de bonnes raisons de « sacrifier » l'uniformité théorique et d'expliquer différemment la motivation liée aux actions morales et la motivation liée aux actions non-morales. Certains estiment que ce sacrifice vaut la peine d'être fait en raison de prises de position substantielles en dehors du débat sur la motivation (sur la nature du normatif ou sur l'essence de la rationalité pratique par exemple)<sup>43</sup>. L'examen des raisons associées à ces prises de positions dépasse évidemment l'objet de la présente recherche. Il suffit seulement de noter que chacune de ces raisons n'a de force que celle de la position sur laquelle elle s'appuie ; position en général plutôt controversée.

Mais, considérons une autre manière par laquelle on pourrait essayer de contrer l'argument de l'uniformité théorique. Il s'agit d'en nier tout simplement la première prémisse. Celle-ci ne fait effectivement pas l'unanimité. Ceux qui sont anti-humiens dans le domaine de la motivation morale peuvent être tentés de la nier et d'endosser une conception anti-humienne de la motivation en général, préservant ainsi

---

43. Jonathan Dancy (Dancy : 2000) et Thomas Scanlon (Scanlon : 1998) sont deux auteurs chez qui l'anti-humianisme semble ainsi fondé sur des prises de position substantielles *en dehors* du débat sur la motivation.

l'uniformité de leur théorie. Nier que la THM est la théorie qui prévaut pour expliquer la motivation dans le cas des actions non-morales, c'est dire que la croyance peut aussi être source de motivation dans les cas *non*-morales. Il se pourrait par exemple que les actions non-morales (ou au moins certaines d'entre elles) soient considérées « *sub specie boni* » par l'agent. En pareil cas, l'agent se verrait motivé par une *croyance* à l'effet que l'action envisagée est une *bonne* action – ou à l'effet que cette action peut procurer une chose *bonne*. Si cela était possible, il n'y aurait alors plus d'avantage à préférer la THM pour expliquer les cas moraux. À vrai dire, si une telle chose était possible et qu'une explication anti-humienne de la motivation était disponible (et meilleure) pour toutes les actions non-morales, alors l'argument de l'uniformité théorique pourrait se renverser complètement : ce serait alors le humien qui romprait l'uniformité de la théorie de la motivation en invoquant une explication des cas moraux différente de l'explication donnée dans les cas non-morales. Il est donc nécessaire que nous nous attardions quelques instants sur la thèse d'une motivation *sub specie boni* dans les cas d'actions non-morales.

### **Les actions non morales : motivation *sub specie boni* ?**

Il faut préciser d'emblée que la possibilité d'une motivation *sub specie boni* dans le cas des actions *non*-morales est en soi l'objet d'un véritable débat. Ainsi, le traitement que je ferai de cette question sera nécessairement superficiel. Il me semble toutefois que certaines considérations plutôt simples s'avèrent déterminantes quant à l'issue de

ce débat. Nous nous en tiendrons donc essentiellement à celles-ci.

Il faut aussi souligner que, sous une forme différente, la thèse de la motivation *sub specie boni*, est habituellement défendue par les théoriciens *humien*s. La raison de cela est que, associée à la thèse de l'internalisme du jugement moral (c'est-à-dire la thèse qui veut qu'il existe un lien interne ou nécessaire entre les jugements moraux et la motivation), la THM conduit au non-cognitivism e sémantique. Nous avons vu en introduction que le non-cognitivism e est une conception de la signification des jugements moraux ; conception selon laquelle les jugements moraux n'ont pas pour fonction de décrire les faits (moraux), mais seulement d'exprimer un désir de l'agent, une attitude d'approbation, etc. Les théories non-cognitivistes tentent, d'une façon ou d'une autre, de rendre compte de la forme *apparente* des jugements moraux (leur aspect « objectif » et « réaliste ») en invoquant une forme plus profonde explicable en termes d'attitudes conatives de l'agent. Ces théories sont qualifiées de « projectivistes » : l'agent projette ses propres attitudes sur les faits du monde et considère que ces faits ont réellement la propriété d'être bons, mauvais, justes, etc. – alors que ces faits sont en eux-mêmes dépourvus de telles propriétés.

Mais, comme le remarque Shafer-Landau, si la thèse de la motivation *sub specie boni* va habituellement de paire avec une conception *humienne* de la motivation, rien n'empêche que cette thèse soit reprise et défendue également par les anti-*humien*s<sup>44</sup>. Nous avons vu jusqu'à maintenant l'anti-*humien* affirmer que la motivation d'un agent dans le cas d'actions *morales* résulte essentiellement d'une croyance à l'effet qu'accomplir cette action est quelque chose de *bon* (ou quelque chose de *juste*, ou

---

44.*Op. cit.*, p. 133.

d'*admirable*, etc.). Mais ce qui doit retenir notre attention ici, c'est la possibilité que l'anti-humien puisse affirmer la même chose dans le domaine *non-moral*<sup>45</sup>.

Évidemment, personne ne soutient que toutes les croyances quelles qu'elles soient peuvent motiver. En effet, personne ne défend l'idée que des croyances telles la croyance *que*  $15 + 17 = 32$  ou la croyance *que Londres est en Angleterre* ont quelque pouvoir motivationnel que ce soit. Outre les croyances liées aux actions morales, les croyances dites *évaluatives* (qui portent sur la valeur d'une action ou d'un état de choses), se voient aussi parfois attribuer le pouvoir de motiver par les anti-humiens. Il est par exemple possible de croire *que ce marteau est un bon marteau* ou *que cet employé est un mauvais employé*. Pareillement pour les actions : on peut croire *qu'il est bon de se protéger du froid en fermant la fenêtre* ou croire *qu'il est mauvais de ne pas préserver sa santé en mangeant beaucoup de légumes* (sans que le contenu de ces croyances ne soit considéré du point de vue moral. Disons qu'il est ici considéré du point de vue prudentiel).

Je crois que la question la plus déterminante pour le problème de la motivation *sub specie boni* dans le cas d'actions non-morales est la suivante : est-ce qu'une explication en termes de croyance au caractère bon de l'action (ou de la chose visée par celle-ci) est disponible pour *toutes* les actions non-morales ? Si l'on parvient à montrer qu'une action au moins n'implique aucune croyance de cette espèce, et puisque l'on sait que la THM explique la motivation pour les toutes les autres actions au moins aussi bien que ses rivales anti-humiennes, alors il devient légitime d'affirmer

---

45.L'enjeu étant, comme nous l'avons vu à la section précédente, que s'il s'avérait que la motivation puisse résulter d'une croyance dans certains cas non-moraux, alors il n'y aurait aucune prévalence de la THM dans les cas non-moraux et, par conséquent, aucune base sur laquelle appuyer le principe d'économie théorique supposé favoriser la THM dans le débat sur la motivation.

la nécessité de la THM (pour expliquer cette action particulière). Dès lors, une généralisation de la THM s'avère la manœuvre la plus parcimonieuse du point de vue théorique.

Nous sommes habituellement prêts à reconnaître que, bien qu'ils ne possèdent pas de croyances évaluatives, les jeunes enfants et les animaux peuvent être motivés à accomplir d'authentiques actions. Un animal ou un enfant en bas âge ne possède sûrement pas le concept de *bon* et ne peut donc pas croire *qu'une chose est bonne*. Si on peut attribuer des croyances à ces êtres, ce ne peut être que des croyances instrumentales. Leur motivation ne découle donc pas d'une croyance puisque, d'une part, personne ne croit que les croyances non-évaluatives (ou non-normatives, ou non-morales, etc.) sont en mesure de motiver d'elles-mêmes et que, d'autre part, aucune croyance évaluative n'est présente dans l'esprit de ces êtres<sup>46</sup>.

Je crois aussi que, dans le monde des humains en pleine possession de leurs capacités conceptuelles, certaines actions simples ne peuvent pas recevoir une explication en termes de croyance au caractère bon de l'action ou de la chose que celle-ci est supposée procurer. Quels exemples a-t-on de cela ? La motivation liée au plaisir et à la douleur corporels. Prenons la motivation à faire cesser la douleur associée à la faim. Je n'ai pas mangé depuis plusieurs heures et ma faim ma cause une vive douleur. Je trouve une pomme sur mon bureau. Je passe à l'action et la mange. Mon action, celle de manger la pomme, est certainement intentionnelle et motivée. Est-il alors

---

46. À l'objection selon laquelle ces êtres pourraient quand même « percevoir » le caractère bon d'un objet de quelque manière, je répondrai que, depuis les études réalisées à la fin des années cinquante par Huber et Wiesel sur les capacités discriminatrices des neurones du cortex visuel, la perception paraît difficilement pouvoir être détachée d'un certain travail *conceptuel* effectué par le système nerveux (pour une revue de ces études, voir Huber et Wiesel 2005).

possible, d'expliquer ce cas précis en invoquant une croyance que j'aurais à l'effet que mon action (ou ce que celle-ci permet d'atteindre) est bonne sous un aspect ou un autre ? Essayons de voir de quoi pourrait avoir l'air une telle explication.

Sous quel aspect l'action de manger cette pomme me paraît-elle bonne ? Une réponse possible est alors que l'action de manger la pomme me paraît bonne *en tant que moyen de soulager ma faim* ; autrement dit, je crois qu'accomplir cette action est un bon moyen de soulager ma faim. Derrière la motivation à manger la pomme, se trouvait donc la motivation à soulager ma faim. En supposant que mon action de manger la pomme puisse être redécrite comme l'action de soulager ma faim (et que, puisque mon action de manger la pomme était intentionnelle, j'étais donc aussi motivé à soulager ma faim lorsque j'ai mangé la pomme), on peut demander en quoi l'action de soulager ma faim me paraît-elle bonne ? Encore une fois, on peut répondre que l'action de soulager ma faim me paraît bonne *en tant que moyen de soulager la douleur causée par celle-ci*. À nouveau, une motivation se cachait derrière une autre ; celle de faire cesser la douleur derrière celle de faire cesser la faim. Supposons encore que mon action de soulager ma faim puisse être redécrite comme l'action de soulager ma douleur (et que, puisque mon action de soulager ma faim était intentionnelle, j'étais donc motivé à soulager ma douleur lorsque j'ai mangé la pomme). Il me semble alors qu'on ne peut plus trouver un aspect sous lequel l'action paraît bonne à l'agent : l'agent ressent la douleur et se voit *immédiatement* motivé à la faire cesser, c'est-à-dire sans que n'intervienne aucun jugement quant à la valeur de l'action de retirer sa main.

Que pourrait-on vouloir dire en affirmant que la perspective de soulager sa douleur apparaît à l'agent comme étant une « bonne chose » et que c'est cela qui le motive ?

Une motivation aussi fondamentale ne peut vraisemblablement découler que de dispositions liées à la physiologie et la psychologie humaine : dans certaines conditions, l'être humain est ainsi fait qu'il devient motivé à accomplir certaines actions. En pareils cas, aucun jugement évaluatif n'intervient dans le schéma expliquant l'action de l'agent.

Peut-être voudra-t-on objecter que même les désirs n'interviennent pas dans le schéma expliquant l'action de l'agent en pareille situation : après avoir ressenti la douleur, l'agent passerait directement de la sensation douloureuse à la motivation, sans la médiation d'un désir. Mais y a-t-il d'abord vraiment lieu de distinguer le désir de l'état de motivation que celui-ci est supposé engendrer ? La réponse à cette question est un problème en soi et dépend de la manière dont on caractérise ce en quoi consiste le désir et ce en quoi consiste la motivation<sup>47</sup>. Remarquons seulement que ce qui sépare humiens et anti-humiens c'est moins la question de la nécessité de la présence d'un désir pour qu'il y ait motivation, que la question de la possibilité pour une croyance de motiver d'elle-même un agent. C'est que les humiens semblent être en mesure d'assouplir leur conception du désir de façon à pouvoir inclure sous sa définition n'importe quel état mental intrinsèquement motivationnel *en autant que celui-ci ne soit pas une croyance*<sup>48</sup>. Peu importe alors s'il y a ou non quelque chose

---

47. Au sein même de la conception humienne de la motivation, il existe de nombreuses manières de définir le désir. Pensons seulement à l'écart qui peut exister entre les conceptions qui confèrent au désir une nature normative et celles qui lui confèrent une nature non-normative. Parmi ces dernières, Smith distingue encore entre les conceptions phénoménologique et les dispositionnaliste (1994 : 104-111). Denis Stampe (1987), pour sa part, propose qu'un désir consiste en un état perceptuel... Bref, caractériser adéquatement le désir constitue un problème à part entière.

48. Tout état mental intrinsèquement motivationnel « qui n'est pas une croyance »... ou dont les caractéristiques essentielles s'apparenteraient trop à celles de la croyance. Comme exemple d'attitude propositionnelle qui pourrait être telle, on pense entre autres à l'attitude qui consiste à *tenir pour hypothèse* [« to hypothetize »] un objet propositionnel donné. À ce sujet, voir aussi ce que je dis à la section 3.2 relativement à la nature non-normative de l'état mental de croyance.

entre la perception de la douleur et la motivation ; quelque chose à quoi on accorderait le nom de « désir » plutôt que de l'accorder à la motivation elle-même – quoi que cette dernière puisse être ainsi séparée de ce qui serait alors le véritable désir. Ce qui caractérise la position humienne, encore une fois, c'est avant tout le refus de concevoir la croyance comme un état mental pouvant de lui-même motiver un agent.

Un agent peut, disions-nous, être motivé sans avoir de croyance évaluative particulière au sujet de son action. En pareille situation, l'état mental de l'agent (l'état motivationnel) est un état dépourvu de fonction représentative (du moins, dépourvu de la fonction de représenter la valeur de l'action ou la valeur de ce que celle-ci permet d'atteindre). Or une croyance évaluative possède *essentiellement* cette fonction représentative. Lorsqu'un agent est ainsi motivé, il est donc nécessaire que son état ne soit pas celui de la croyance. Il existe donc des actions qui ne peuvent être expliquées (du point de vue motivationnel) qu'en recourant à un désir ou une autre « pro-attitude » dépourvue de la propriété de représenter la valeur d'une action ou d'un état de choses.

Gardons en tête que l'anti-humien n'a pas en sa faveur d'exemples d'action qui nécessiterait une explication formulée uniquement en termes de croyances (parce qu'aucune explication humienne ne serait disponible pour en rendre compte). Autrement dit, toute action pour laquelle l'anti-humien possède une explication peut aussi être expliquée en termes strictement humiens. Cela exige parfois d'en appeler à des désirs qui paraissent quelque peu ad hoc, mais dont l'existence n'est pas pour autant invraisemblable.

Si vraiment les choses sont ainsi, alors il existe une bonne raison d'adopter la

position humienne dans le débat sur la motivation en général. Il arrive en effet qu'un agent soit motivé à accomplir une action non-morale sans avoir de croyance évaluative particulière au sujet de l'action en question (ou de ses conséquences) – le cas des animaux et des jeunes enfants illustre cela de manière particulièrement claire. Une explication de type humien est alors absolument nécessaire puisque ces actions ont pour condition suffisante un état mental qui motive mais qui n'est pas doté de la qualité représentative habituellement associée à ce type croyance. Et puisqu'il n'existe aucune action pour laquelle on ne dispose pas d'une explication humienne au moins aussi satisfaisante que l'explication anti-humienne, alors il s'avère plus parcimonieux de généraliser notre explication (déjà nécessaire) – de ces cas particuliers vers tous les cas non-moraux – que d'attribuer une nouvelle propriété (celle de motiver ou celle de générer un état intrinsèquement motivationnel) aux autres croyances non-morales. Ceci constitue le premier mouvement d'une généralisation qui, en un second mouvement, doit étendre l'explication humienne de façon à inclure aussi les actions morales. Si l'on généralise l'explication humienne de certains cas non-moraux à tous les cas non-moraux et que la THM devient l'explication primée dans le cas de toutes les actions non-morales, alors il est en effet souhaitable (du point de vue de l'économie théorique) d'adopter la THM en tant que théorie *générale* de la motivation.

## **2.4 Conclusion du chapitre**

Résumons notre examen des trois principaux arguments en faveur de la théorie humienne de la motivation. Nous avons vu que l'argument des croyances identiques consiste à inférer l'inertie de la croyance à partir de la prémisse selon laquelle il est nécessaire que deux agents qui possèdent les mêmes croyances soient identiquement motivés et de la prémisse selon laquelle il est de fait possible que deux agents possèdent les mêmes croyances sans être identiquement motivés. L'idée est que si une croyance était suffisante pour motiver, alors son introduction dans l'esprit de l'agent A devrait avoir le même effet (du point de vue motivationnel) que son introduction dans l'esprit de l'agent B. Il semble toutefois que l'argument n'ait rien de décisif puisque l'on peut nier de façon persuasive que deux agents qui partagent les mêmes croyances doivent être identiquement motivés. Il se peut en effet que les croyances soient bel et bien suffisantes pour motiver, mais que deux agents qui partagent les mêmes croyances soient quand même différemment motivés : d'une part, la présence d'autres états mentaux dans l'esprit de chaque agent peut exercer une influence sur le pouvoir motivationnel des ces croyances ; d'autre part, la manière dont chaque agent « perçoit le monde » peut aussi exercer une influence sur la pouvoir motivationnel d'une croyance. La première de ces hypothèses nous rappelle que l'anti-humien soutient seulement que certaines croyances motivent *ceteris paribus*, c'est-à-dire lorsque leur pouvoir n'est pas supplanté (*overridden*) par celui d'autres états mentaux. La deuxième hypothèse met l'accent sur le « cadre conceptuel » (qui n'est pas seulement l'ensemble des croyances d'un agent) dans lequel s'inscrit une croyance ; inscription qui (telle un

*gestalt switch*) change la perception globale de l'agent et fait passer sa croyance d'inerte à potente.

Concernant l'argument de la direction d'ajustement, nous avons vu que celui-ci consiste à dire que désirs et croyances se distinguent par leur « direction d'ajustement », c'est-à-dire par le fait que la croyance est un état dont le contenu doit s'ajuster aux faits du monde, tandis que le désir est un état au contenu duquel les faits du monde doivent s'ajuster. Il s'agit ensuite d'essayer de donner un sens non-métaphorique à la notion de direction d'ajustement. Ces tentatives se résument à la démonstration de l'inertie de la croyance sur la base du fait qu'il est impossible pour un même état d'instancier simultanément des directions d'ajustement opposées (cela exigerait en effet de manifester simultanément des dispositions incompatibles).

Avec l'interprétation dispositionaliste donnée par Smith à la notion de direction d'ajustement, l'idée selon laquelle les faits du monde « doivent s'ajuster » au contenu d'un désir devient l'idée que le désir est un état qui tend à persister à l'occurrence d'une perception de contenu contraire et qui dispose l'agent à réaliser le contenu de son état mental ; l'idée selon laquelle une croyance « doit s'ajuster » aux faits du monde devient l'idée que celle-ci tend à disparaître à l'occurrence d'une perception de contenu contraire. La présence d'un désir s'avère donc nécessaire pour motiver un agent à l'action puisqu'une croyance ayant le pouvoir de motiver est un état qui devrait manifester simultanément une disposition à persister et une disposition à disparaître à l'occasion d'une perception de contenu contraire.

Selon la conception normative de Zangwill, le contenu d'un état *doit* s'ajuster aux faits du monde (ou inversement, les faits du monde s'ajuster au contenu d'un état). La

normativité des états mentaux est issue de leur nature rationnelle. La distinction des désirs et des croyances tient donc au fait qu'une perception peut rationaliser une croyance sans pouvoir rationaliser un désir et qu'un désir peut rationaliser une intention (ou une action) alors qu'une croyance ne le peut pas. Comme Smith, Zangwill pense que les dispositions que devrait manifester une croyance capable de motiver rendent un tel état inconcevable. (Il faut cependant noter que, contrairement à ce que l'on observe chez Smith, les dispositions ne définissent pas les états mentaux auxquels elles sont associées et découlent plutôt de leur nature – intrinsèquement normative puisqu'il est dans l'essence de tel état de rationaliser tel autre état ou d'être rationalisé par tel autre état encore).

Quant à l'argument de l'uniformité théorique, il nous a paru constituer une raison d'une certaine force en faveur de l'adoption de la position humienne dans le débat sur la motivation. (Nous n'avons toutefois pas encore considéré les principales objections adressées à la THM.) Il arrive en effet qu'un agent soit motivé à accomplir une action non-morale sans avoir de croyance évaluative particulière au sujet de cette action ou de ce qui découlera de celle-ci (dans le cas d'actions simples liées à nos dispositions naturelles fondamentales et aussi dans le cas d'actions posées par de jeunes enfants ou des animaux). En pareille situation, la croyance n'est pas une condition nécessaire de la motivation : l'état motivationnel est dépourvu de propriété représentative (du moins en ce qui a trait à la valeur de l'action envisagée). Il doit s'agir d'un désir ou d'une autre attitude purement conative. Il existe donc des actions non-morales qui *nécessitent* une explication humienne. Or, puisque toutes les actions susceptibles d'être expliquées en termes anti-humiens sont aussi susceptibles de l'être en termes

humains, il est plus parcimonieux d'étendre l'explication humaine aux autres cas non-moraux que d'attribuer une nouvelle propriété aux croyances non-morales (la propriété de motiver par elle-même ou de celle de générer un état intrinsèquement motivationnel). Et si l'on généralise ainsi l'explication humaine à tous les cas non-moraux, alors il est souhaitable (pour la même raison) de choisir la théorie humaine en tant que théorie générale de la motivation.

Passons maintenant à l'examen des objections adressées à la THM. Les interprétations de l'argument de la direction d'ajustement discutées ici (celles de Smith et Zangwill) seront confrontées aux deux premières de ces objections : l'objection de la circularité de la caractérisation de la différence entre désirs et croyances et l'objection des croyances irrationnelles. Suivra ensuite ce qui, à mon avis, constitue l'objection la plus intéressante adressée à la THM, à savoir l'objection des états mentaux de double direction d'ajustement.

## **Chapitre 3**

### **Arguments à l'encontre de la**

### **Théorie Humienne de la Motivation**

Nous avons déjà examiné les considérations relatives à la phénoménologie de la motivation morale parfois soulevées à l'encontre de la THM. Celles-ci auraient pu constituer l'« objection de la phénoménologie de la motivation morale » et recevoir une place parmi les autres objections à la THM. (On se rappelle que Shafer-Landau fait de cette objection une pièce centrale de sa défense anti-humienne.) Mais je crois toujours qu'il valait mieux traiter cette objection comme une considération préliminaire à l'examen des arguments pro et anti-humiens<sup>49</sup>. Seuls trois arguments feront donc l'objet du présent chapitre. Les deux premiers arguments sont des

---

49.C'est que l'objection est à mon sens mineure. Rosati a bien fait remarqué que le humien pouvait fournir une explication au moins aussi « directe » et « naturelle » que l'anti-humien pour les cas où l'agent rapporte ne pas avoir perçu la présence de quelque désir que ce soit parmi ses état mentaux au moment d'agir. L'examen précoce des considérations relatives à la phénoménologie de la motivation morale a notamment fourni une occasion de montrer que la THM est au moins superficiellement plausible en montrant comment celle-ci pouvait accommoder certains cas qui, à première vue, semblent difficilement pouvoir recevoir une explication humienne.

objections formulées directement à l'endroit de l'argument de la direction d'ajustement : soit celle consistant à dire que la distinction entre désirs et croyances est circulairement caractérisée et donc fondamentalement viciée ; et celle qui consiste à faire remarquer qu'il existe des croyances irrationnelles et que celles-ci ne répondent pas à la définition de la croyance donnée par la théorie de la direction d'ajustement. Le troisième argument sera celui qui retiendra le plus notre attention ici. Il s'agit de l'argument des états mentaux de double direction d'ajustement. Bien que celui-ci ait aussi à voir avec l'argument de la direction d'ajustement, ce sera surtout son rapport avec la question de l'uniformité théorique qui nous intéressera.

Voyons tout de suite la première objection adressée à l'argument de la direction d'ajustement, à savoir celle qui consiste à dire que la distinction entre désirs et croyances est circulairement caractérisée.

### **3.1 L'argument de la circularité et de la non-informativité**

#### **L'argument**

On se rappelle que Smith propose d'interpréter de manière dispositionaliste l'idée

métaphorique d'une distinction au niveau de la direction d'ajustement des croyances et des désirs :

Grossièrement, une croyance que *p* est un état qui tend à cesser d'exister en présence d'une *perception que non p*, tandis qu'un désir que *p* est un état qui tend à persister, disposant le sujet qui se trouve dans cet état à faire en sorte que *p* soit réalisé<sup>50</sup>.

La formule proposée par Smith soulève le problème de l'interprétation de l'expression « perception que non-*p* ». Il va de soi que la perception en question est un élément central de cette formule et que la distinction entre croyances et désirs en dépend : c'est par les dispositions qu'ils manifestent à l'égard de *cette perception* que désirs et croyances se définissent. Le problème est précisément là.

Lloyd Humberstone (1992 : 63-65) affirme qu'il existe seulement deux lectures possibles de l'élément de définition proposé par Smith<sup>51</sup>. La première lecture, la lecture dite « subdoxastique », attribue au terme « perception » le sens d'*apparence perceptuelle*. La « perception que non-*p* » ne possède alors pas la direction d'ajustement esprit-monde (celle de la croyance). Selon Humberstone, donner ce sens au terme « perception » revient à dire que, *par défaut*, l'agent considère les apparences en question comme représentatives de ce qui est vraiment le cas. Ce n'est qu'à titre

---

50.Smith, 1987, p. 54. Mes italiques.

51.Notons que, bien qu'il rejette l'interprétation dispositionaliste offerte par Smith, Humberstone propose lui aussi une interprétation descriptive de l'idée de direction d'ajustement, au sens où son interprétation rend compte de comment *sont* les choses et non de comment les choses *doivent être*. Humberstone ne fait toutefois pas reposer la distinction entre désirs et croyances sur les conditions de dépendance contrefactuelles des états mentaux concernés relativement une perception de contenu contraire, mais plutôt sur une « intention de contrôle de second ordre » dont la présence ou l'absence détermine le type d'attitude dont il s'agit : « à moins que celui qui se trouve dans un attitude donnée n'ait l'intention de contrôle de second ordre [*background controlling attitude*] que son attitude succède seulement si son contenu propositionnel est vrai, alors l'attitude adoptée n'est pas celle de la croyance » (Humberstone 1992, p. 73).

exceptionnel que l'agent suspend son jugement à l'égard de ces perceptions ou bien qu'il juge que celles-ci ne doivent pas l'inciter à réviser ses croyances. David Copp et David Sobel font le même constat et donnent en exemple la situation d'un conducteur qui voit ce qui ressemble à une mare d'eau au loin sur la chaussée alors qu'il circule sur une chaussée pourtant chaude et sèche. Si le conducteur est expérimenté, sa croyance à l'effet que la route est sèche ne tendra aucunement à disparaître à l'occasion d'une telle perception. La lecture subdoxastique doit donc être écartée puisque, comme le disent Copp et Sobel, « il semble y avoir un large éventail de cas pour lesquels la théorie donne des résultats incorrects. »<sup>52</sup>

Selon la deuxième lecture, la lecture dite « doxastique », l'élément de définition « la perception que non-p » renvoie à une croyance de l'agent dont le contenu serait *non-p*. Supposons donc que « la perception de contenu non-p » signifie *la croyance de contenu non-p*. De l'avis d'Humberstone, « nous sommes ici en train d'expliquer la direction d'ajustement thétique [esprit-monde] en référant à des états ayant la direction thétique »<sup>53</sup> ; autrement dit, la direction d'ajustement de la croyance est présupposée dans sa propre définition. Mais il est à noter que le problème sur lequel Humberstone veut attirer l'attention est le fait qu'une telle caractérisation est *non-informative*. Selon lui, caractériser ainsi la distinction au niveau des directions d'ajustement du désir et de la croyance, c'est exactement comme suggérer « qu'il existe la suivante asymétrie profonde entre les hommes et les femmes en ce qui a trait à la sexualité : tandis qu'un

---

52.Copp et Sobel, 2001, p. 49.

53.Humberstone, 1992, p. 63. Humberstone forge les termes « thétique » et « télique » afin de disposer de mots simples pour référer aux différentes directions d'ajustement ; mots qui ne seraient pas lestés du poids des associations qui viennent avec les termes de la distinction « théorique/pratique » ou ceux de la distinction « cognitif/conatif ».

homme hétérosexuel ne sera *pas* sexuellement attiré par les hommes, une femme hétérosexuelle *le sera*. »<sup>54</sup> Évidemment, le problème en question vient de ce que la distinction croyance-désir est caractérisée de manière circulaire. Et la critique de Copp et Sobel, quant à elle, pointe directement la circularité de l'énoncé définissant cette distinction (ou définissant la nature des états mentaux) – énoncé compris selon la lecture doxastique rappelons-le<sup>55</sup>.

Donc, si vraiment la lecture doxastique est la bonne et que la distinction entre désir et croyance formulée par Smith fait ainsi vicieusement appel à ce qu'elle doit définir, alors l'argument de la direction d'ajustement est privé d'une de ses interprétations majeures.

De plus, il semble que l'interprétation normative de Zangwill soit sujette à la même objection. C'est parce que cette interprétation fait aussi référence de manière essentielle à la perception problématique :

L'asymétrie que nous considérons, [...] est que le désir que *p*, par lui-même, rationalise l'intention de faire en sorte que *p*, mais la croyance que *p* ne rationalise pas cette intention. Et l'*expérience perceptuelle que p*, par elle-même, rationalise la croyance que *p*, mais ne rationalise pas le désir que *p*<sup>56</sup>.

---

54.Humberstone, 1992, p. 64.

55.Humberstone affirme que le souci (avec l'interprétation de Smith) « n'est pas que quelque prétendue analyse du concept de croyance échoue parce qu'elle emploie, de façon déguisée, ce concept même » (p. 64). Cf. Copp et Sobel : « Ce qui n'ennuie pas Humberstone nous ennuie nous, puisqu'il nous apparaît que Smith désirait vraiment offrir une telle analyse » (p. 46). Humberstone semble penser que le projet de Smith est seulement d'offrir un critère qui permettrait de distinguer croyances et désirs, *sans nécessairement caractériser la nature de ceux-ci*. Il me semble qu'il n'y a pas d'incohérence dans cette idée. Mais, avec la tentative de Smith, cela ne fonctionne tout simplement pas (pour la raison mise en évidence par Humberstone).

56.Zangwill, 2005, p. 195. Mes italiques.

La situation reste la même. D'une part, la formule peut recevoir une lecture subdoxastique et l'expression problématique signifiera quelque chose comme « apparence perceptuelle ». Il n'y aura alors pas de circularité nocive, mais le test de la direction d'ajustement donnera de mauvais résultats pour de nombreux cas. (En reprenant l'exemple précédant, la perception par un conducteur expérimenté de ce qui semble être une mare d'eau au loin sur la chaussée chaude et sèche *ne rationalise pas* la croyance chez lui à l'effet que la chaussée est trempée.) Le critère ne peut fonctionner que dans le cas où l'agent tient d'emblée ses perceptions pour véridiques au lieu de les considérer comme fausses ou de suspendre son jugement à leur sujet. D'autre part, la formule peut recevoir une lecture doxastique ; mais alors se pose le problème de circularité dont nous venons de discuter.

L'objection de la non-informativité et de la circularité est, me semble-t-il, d'une grande force à l'égard de l'argument de la direction d'ajustement. De façon très simple, elle met à jour un vice majeur des tentatives faites pour interpréter de l'idée métaphorique de direction d'ajustement. Je ne suis pas certain qu'il soit possible de modifier ou de lire autrement les interprétations de Smith et Zangwill de manière à pouvoir répondre à cette objection. Nous allons voir à l'instant que, de toute façon, un autre problème grave (peut-être impossible à surmonter dans le cas de Smith) se présente pour ces interprétations, à savoir celui de l'existence des croyances irrationnelles.

### 3.2 L'existence des croyances irrationnelles

#### L'argument

Au-delà des difficultés que nous venons d'examiner, les interprétations dispositionaliste et normative de l'argument de la direction d'ajustement semblent aux prises avec une difficulté plus grave encore. Il semble en effet que ces interprétations soient incapables d'accommoder une certaine intuition que nous avons au sujet de ce qui compte ou ne compte pas comme une croyance. Si, comme je le crois, cette intuition est juste, alors cela montre que les conditions que doit remplir un état mental pour être considéré (selon ces interprétations) comme une croyance ne sont tout simplement pas nécessaires.

L'intuition en question est celle qui nous fait considérer comme une croyance l'état mental dans lequel se trouve parfois un individu en proie à une forme grave d'irrationalité. Nous pensons habituellement que l'état en question est une croyance, même si cet état ne manifeste aucune tendance à disparaître à l'occurrence d'une perception de contenu contraire. Je crois que ce qu'on appelle en anglais *wishful*

*thinking* constitue un exemple d'état mental que l'on considère habituellement comme une croyance bien qu'il ne présente pas la tendance à disparaître prétendument caractéristique des croyances. Les croyances d'un agent en proie au *wishful thinking* sont, pour ainsi dire, entretenues sur la base de désirs présents dans son esprit. Tant que ces désirs persistent, le rapport des croyances de l'agent aux faits du monde (et donc aux perceptions de contenu contraire) revêt assez peu d'importance.

Si jamais on voulait objecter que les cas de *wishful thinking* ne mettent pas en jeu de « véritables » croyances (en prétextant disons que l'agent en pareille situation n'est pas vraiment sincère), alors il faudrait considérer un exemple plus fort. Imaginons quelqu'un en proie à un délire paranoïaque qui affirme (sincèrement) être l'objet d'un complot. Malgré toutes les preuves qu'on peut lui apporter à l'effet qu'il n'en est rien (« perceptions » de contenu contraire), cet individu reste toujours dans le même état mental. Il paraît alors assez naturel de dire que l'état mental en question en est un de croyance, et ce même si cet état ne tend pas à disparaître quand surviennent des perceptions de contenu contraire. Il semble en être de même pour les croyances caractéristiques de la plupart des formes de délires, obsessions, et autres représentations distordues de la réalité. Ceux qui sont en proie à celles-ci sont habituellement considérés comme *croyant* réellement les propositions correspondant à leur représentation des choses.

Ainsi, les croyances des individus atteints de formes d'irrationalité graves constituent des contre-exemples à la conception dispositionaliste des croyances (et donc aussi à la conception dispositionaliste des états mentaux en général). Pour qu'un état mental soit considéré comme une croyance, il n'est simplement *pas nécessaire*

que celui-ci présente une tendance à disparaître à l'occurrence d'une perception de contenu contraire à son propre contenu.

Peut-être l'existence des croyances irrationnelles pose-t-elle problème de manière plus évidente pour l'interprétation dispositionaliste que pour l'interprétation normative, mais cette dernière semble elle aussi souffrir du fait de leur existence. C'est que notre intuition au sujet des croyances irrationnelles (à savoir que celles-ci comptent pour d'authentiques croyances) peut nous inciter à réviser les conditions que devrait satisfaire un état mental pour être considéré comme une croyance ; et mettre en doute le fait que la sujétion à des normes de rationalité soit une de ces conditions.

On pourrait en effet être tenté de nier le statut fondamental des normes de rationalisation (qui, selon Zangwill, définissent la nature de chaque état mental) et poser que celles-ci doivent *découler* d'une (ou plusieurs) propriété non-normative qui, elle, est *constitutive* de l'essence d'un état mental. Les croyances irrationnelles, qui sont peut-être les plus à la limite de ne pas être considérées comme d'authentiques croyances, semblent quand même pouvoir être saisies par une définition strictement non-normative. Copp et Sobel par exemple, suggèrent qu'une croyance (rationnelle ou non) pourrait se définir comme l'état mental dans lequel se trouve un agent qui tient un contenu propositionnel donné pour vrai<sup>57</sup>. Que l'agent tienne ainsi un contenu pour vrai, disons *p*, cela ne dépend aucunement du caractère rationnel (ou non) du fait de

---

57.« [L]'idée que les croyances, contrairement aux désirs, sont rationalisées par des expériences perceptuelles de contenu approprié présuppose qu'il y a quelque différence non-normative entre les croyances et les désirs. Par exemple, on pourrait dire qu'avoir une croyance *consiste à tenir un contenu [propositionnel] pour vrai*, et les expériences perceptuelles peuvent être des preuves en faveur de la vérité de leurs contenus. » (Copp et Sobel, p. 52. Mes italiques.) Cette thèse n'est évidemment pas équivalente à la thèse dénoncée (avec raison) par Zangwill ; celle voulant que les croyances soient des états qui peuvent être vrais ou faux – tandis que les désirs seraient des états qui peuvent « être satisfaits » ou non (1998, p. 173).

tenir  $p$  pour vrai lorsque cet agent perçoit que  $p$ . Autrement dit, le statut rationnel ou irrationnel de ce fait (que l'agent tienne  $p$  pour vrai lorsqu'il perçoit que  $p$ ) pourrait varier (à supposer que cela soit possible) sans que cela n'affecte le fait que l'agent tienne  $p$  pour représentation du monde tel qu'il est. Les relations de rationalisation ne détermineraient donc pas l'essence de la croyance. Bien sûr les croyances, même irrationnelles, sont *sujettes* à des normes de rationalité. Mais ces dernières ne sont pas *constitutives* de leur essence.

Soyons bien clairs relativement à la nature exacte du problème soulevé par l'existence des croyances irrationnelles pour l'interprétation de Zangwill. Ce n'est très certainement pas que cette interprétation ne peut pas expliquer les croyances en questions ou qu'elle définirait la croyance de manière à ce que les cas de croyances irrationnelles ne comptent pas comme d'authentiques croyances. Il n'y a aucune difficulté à ces deux niveaux. La croyance que  $p$  est l'état mental dans lequel il est rationnel de se trouver lorsqu'on perçoit que  $p$ , mais cette croyance ne manifestera pas nécessairement une tendance à disparaître à l'occasion de la perception que non- $p$ . Et, si c'est le cas qu'elle ne manifeste pas cette tendance et qu'elle est irrationnelle, il n'en demeure pas moins qu'elle est toujours jugée relativement au canon de la raison théorique (c'est à ce compte qu'elle est « irrationnelle »), et en tant qu'il est dans sa nature d'être jugée comme telle, elle remplit la condition nécessaire pour être considérée comme une croyance selon Zangwill. Le problème vient plutôt de ce que la thèse selon laquelle la nature des états mentaux est normative est une thèse problématique en soi : on pourrait soutenir que la différence normative entre désirs et croyances découle elle-même de quelque différence non-normative au sein de leur

nature. Notre manière de concevoir les croyances irrationnelles irait dans ce sens puisque cela semble une caractérisation suffisante de celles-ci que d'affirmer qu'elles sont l'état dans lequel se trouve un agent lorsque ce dernier tient le contenu propositionnel de l'état en question pour vrai.

À ce point, les deux principales interprétations de l'idée de direction d'ajustement paraissent souffrir de graves problèmes. L'objection qui suit sera, en un premier temps, dirigée elle aussi contre l'argument de la direction d'ajustement. Il s'agit de l'objection voulant que certains états mentaux possèdent une double direction d'ajustement relativement à deux contenus différents. Mais, en un second temps, cette objection prendra aussi pour cible l'argument de l'uniformité théorique. Ceci sera d'un grand intérêt puisque, au vu des résultats de la critique adressée à l'argument de la direction d'ajustement, l'argument de l'uniformité théorique semble être le seul à pouvoir offrir encore quelque appui à la théorie humienne de la motivation<sup>58</sup>.

### **3.3 Argument des états mentaux de double direction d'ajustement**

À l'argument selon lequel croyances et désirs se distinguent par leur direction

---

58. Rappelons en effet que, au chapitre consacré aux arguments en faveur de la THM, nous avons déjà conclu que l'anti-humien dispose de réponses convaincantes à l'argument des croyances identiques (voir pp. 29-34). Si l'argument de la direction d'ajustement ne fonctionne pas, alors l'argument de l'uniformité théorique est vraiment le seul des arguments considérés ici qui puisse (peut-être) offrir un appui à la THM.

d'ajustement et selon lequel seul le désir possède la direction d'ajustement appropriée pour rendre compte du fait de poursuivre un but (i.e. d'être motivé), nous verrons qu'il est possible de répondre en soutenant que certains états mentaux présentent simultanément les deux directions d'ajustement sans que cela n'implique pour eux de manifester des dispositions incompatibles (la tendance à persister à l'occurrence d'une perception de contenu contraire et la tendance à disparaître à l'occurrence de cette même perception). L'idée sera de faire porter chaque direction sur un contenu propositionnel différent. Les considérations relatives à ces états de double direction d'ajustement nous mèneront ensuite à examiner la valeur de l'argument de l'uniformité théorique. C'est que, à première vue au moins, certaines découvertes du côté de la neurophysiologie de la perception et de la motivation semblent indiquer qu'il existe une bonne raison de vouloir rompre l'uniformité théorique et admettre que certaines croyances au moins sont capables de motiver par elles-mêmes.

### **États mentaux de double direction d'ajustement**

Faisons donc comme si les difficultés précédemment exposées n'étaient pas décisives quant au succès de l'argument de la direction d'ajustement. Dans son interprétation dispositionaliste, rappelons-le, l'argument de la direction d'ajustement montre que l'identité de l'état motivationnel ne peut être celle de la croyance : étant donné que, d'une part, l'état motivationnel tend à persister à l'occurrence de

perceptions de contenu contraire, et que, d'autre part, (selon cette interprétation) une croyance tend à disparaître à l'occurrence des mêmes perceptions ; l'état motivationnel devrait donc manifester des dispositions incompatibles et cela est bien sûr impossible. Philip Pettit a toutefois fait remarquer que le problème ne survient que si l'on présuppose que l'état en question a deux directions d'ajustement opposées *en rapport avec un contenu unique*<sup>59</sup>. C'est n'est en effet que si la « croyance motivationnelle » se comporte pour une part comme une croyance *de contenu p* et, pour une autre part, comme un désir *de contenu p* qu'elle devra manifester des dispositions incompatibles à l'occasion d'une perception de contenu non-p. Mais qu'est-ce qui empêche de concevoir qu'il existe des états mentaux ayant deux directions d'ajustement *relativement à deux contenus propositionnels distincts* ? Little exprime la chose de la manière suivante :

Certains semblent présupposer que là où il y a un état mental il y a un contenu propositionnel ; ainsi lorsque confrontés à la possibilité d'un état avec une double direction d'ajustement, ils se demandent quel pourrait bien être « son » contenu propositionnel. Mais un état mental avec une double direction d'ajustement n'a pas qu'un contenu propositionnel, pas plus qu'un tel état n'implique qu'une seule attitude propositionnelle. Il s'agit d'un état possédant deux propriétés *complexes* : c'est une attitude de croyance dirigée vers une proposition, et c'est une attitude de désir envers une autre. (Little, 1997 : 64)

Nous avons déjà rencontré un des états mentaux qui semblent les plus susceptibles de se voir attribuer une double direction d'ajustement relativement à deux contenus propositionnels. Il s'agit de la croyance évaluative. Celle-ci aurait la direction

---

59.Pettit, 1987, pp. 530-531.

d'ajustement de la croyance relativement au contenu « x est bon » ; et la direction d'ajustement du désir relativement à au contenu « x ». Un tel état présenterait donc simultanément les caractéristiques du désir et de la croyance, et ce *en tant qu'état unique*, c'est-à-dire que la croyance évaluative ainsi conçue ne consiste pas en une simple juxtaposition de deux états préexistants (le désir *que x* et la croyance *que x est bon*). Chaque direction d'ajustement serait en fait associée à un « aspect » de cet état (et non à un état mental préexistant à l'état motivationnel). Le terme souvent utilisé pour désigner un tel état est celui de « croisir » (formé à partir du « croi » de *croyance* et du « sir » de *désir*) ou, en anglais, celui « besire »<sup>60</sup>.

La perception dont le contenu est « non-x » doit seulement être mise en relation avec l'aspect du croisir qui s'apparente au désir et dont le contenu est « x ». Ainsi, l'état mental de celui qui perçoit que « non-x » n'a pas à manifester des dispositions incompatibles : l'aspect de cet état qui s'apparente au désir persiste à l'occurrence de la perception et l'aspect de ce même état qui s'apparente à la croyance persiste à cette perception – lui dont le contenu n'est pas « x » mais plutôt « x est bon ». Un croisir serait donc un type de croyance qui peut prendre « x » pour contenu, ne pas tendre à

---

60. Le terme « besire » fût introduit par J.E.J. Altham (1986). Pour sa traduction par le terme « croisir », voir Virvidakis, 1996, p. 191. Aux lecteurs qui s'inquiéteraient de l'absence d'attention portée jusqu'ici aux « desire-entailing-beliefs » (celles-ci ayant à peine été mentionnées au précédent chapitre), je dirai seulement que le traitement qu'il convient d'en donner est, à mon sens, le même que celui qui s'applique aux croisirs ou, plus généralement, à toute croyance prétendument dotée du pouvoir de motiver par elle-même. Une desire-entailing-belief est une croyance dont la présence dans l'esprit d'un agent garantit que cet agent aura aussi un désir qui le motivera conformément au contenu de la croyance en question. Little (1997, p. 70) souligne le fait que cette thèse peut prendre deux sens bien distincts. Cela peut vouloir dire que la croyance de l'agent a pour condition un certain désir ; auquel cas la croyance n'aura aucune force motivationnelle intrinsèque et le humien rien à objecter. Mais cela peut aussi vouloir dire que la présence du désir en question est expliquée par celle de la croyance ; ce qui est l'équivalent d'affirmer que la croyance est dotée de la propriété de *générer d'elle-même un état mental intrinsèquement motivationnel*. Pour cette croyance comme pour la croyance dotée du pouvoir de motiver par elle-même, le humien fera valoir qu'il vaut mieux s'abstenir d'invoquer de nouvelles propriétés et généraliser l'explication humienne.

disparaître à l'occurrence d'une perception de contenu « non-x » et, ce faisant, ne pas contredire la définition de la croyance selon laquelle une croyance est un état qui tend à disparaître à l'occurrence d'une perception de contenu contraire à son propre contenu (puisqu', encore une fois, le contenu propositionnel lié à l'aspect proprement cognitif de cette croyance est « x est bon » et non pas « non-x »).

Contre cette idée d'un état mental de double direction d'ajustement, Smith fait valoir que la possibilité d'états mentaux de double direction d'ajustement présuppose l'existence d'un lien beaucoup trop fort entre croyance et motivation (entre les « aspects » cognitif et conatif de l'état motivationnel, dirait-on plutôt avec les tenants de la thèse du croisir)<sup>61</sup>. Certains cas de dépression, de fatigue excessive, de faiblesse de la volonté, etc. constituent des cas évidents de rupture du lien en question. Dans de tels cas, le sujet *croit* que l'action considérée est ou juste, ou bonne, ou obligatoire, etc., mais il *n'est pas motivé* à accomplir cette action (en tout cas pas *suffisamment* motivé pour que cela se traduise en action). Nous avons déjà vu que cette objection était sans force, puisqu'il existe au moins deux manières plausibles d'expliquer en termes anti-humiens comment le pouvoir motivationnel de la croyance, bien que réel, peut ne pas se traduire en action : une croyance peut ne pas motiver (ou se traduire en action) si d'autres états mentaux interfèrent avec elle et inhibent son pouvoir motivationnel ; une croyance peut aussi ne pas motiver (ou se traduire en action) si elle ne se trouve pas subsumée sous le « cadre conceptuel » approprié. Il s'agit-là des réponses déjà offertes à l'argument des croyances identiques. Celles-ci expliquent comment, dans certaines conditions, le pouvoir motivationnel du croisir peut ne pas se

---

61.Smith, 1994, p. 120.

manifester, et donc comment une thèse telle celle du croisir n'implique rien d'in vraisemblable eu égard au lien entre la croyance et le désir (ou entre les aspects cognitif et conatif de l'état motivationnel).

On pourrait toutefois reprocher au théoricien anti-humien de proposer (avec sa thèse du croisir) une solution purement ad hoc. Voulant éviter le problème soulevé par l'argument de la direction d'ajustement, celui-ci propose *sans justification* d'individualiser les états mentaux d'une manière plutôt inhabituelle (sur la base de contenus propositionnels multiples). Comme le fait remarquer Shafer-Landau, si, en dehors des cas mettant en scène de soi-disant croisirs, on individualise tous les états mentaux sur la base de contenus uniques, alors on ne voit pas très bien pourquoi il faudrait faire une exception ici<sup>62</sup>. Et c'est sans doute un présupposé plausible que de penser que les états mentaux doivent être individualisés à partir de contenus propositionnels uniques.

Quoi qu'il en soit du nombre de contenus, tant que l'idée d'un état mental possédant les deux directions d'ajustement demeure une hypothèse et ne s'est pas montrée nécessaire pour expliquer certains cas de motivation, le humien pourra faire valoir l'argument de l'uniformité théorique et ainsi nous convaincre d'adopter la THM en tant que théorie générale de la motivation. Mais nous allons voir à l'instant que l'anti-humien peut peut-être s'appuyer sur une découverte empirique pour montrer qu'un état mental de double direction d'ajustement existe de fait. La découverte en question est celle de l'existence d'un groupe de neurones dans le cerveau de certains mammifères ; neurones qui sont dotés de propriétés particulières et qui pourraient peut-être constituer le substrat de ces fameuses croyances dotées du pouvoir de motiver par

---

62. Shafer-Landau, p. 137.

elles-mêmes (ou du pouvoir de générer par elles-mêmes un état intrinsèquement motivationnel). S'il s'avérait que ces neurones accomplissent véritablement (en tant que substrat bien sûr<sup>63</sup>) les fonctions du désir et de la croyance, alors cela constituerait une bonne raison de rompre l'uniformité théorique de notre théorie générale de la motivation : s'il y a des cas où il est clair que la motivation surgit d'un état unique présentant simultanément les caractéristiques de la croyance et du désir, alors il n'y aura pas d'économie à réaliser en généralisant l'explication humienne (qui s'était montrée nécessaire pour expliquer certaines actions simples, notamment celles des jeunes enfants et des animaux et celles liées à nos dispositions physiologiques et psychologiques les plus fondamentales : se nourrir, avoir des rapports sexuels, éviter la douleur corporelle, etc.). Il faut donc que nous voyions ce que la neurophysiologie peut nous apprendre au sujet de l'existence d'états mentaux qui présentent à la fois les caractéristiques du désir et celles de la croyance. Les dernières pages de ce chapitre seront dédiées à cette tâche.

### **Neurones miroirs et intentionnalité**

Dans un contexte autre que celui de Pettit, Ruth Gareth Millikan soulève elle aussi la possibilité de l'existence d'états mentaux de double direction d'ajustement : selon elle, il y a de bonnes raisons de penser qu'il existe des états mentaux qui, selon sa

63.C'est dire que si le désir et la croyance possèdent des caractéristiques mentales essentielles, les neurones en question en constitueront la base matérielle de survenance.

terminologie, sont à la fois « descriptifs » et « directifs »<sup>64</sup>. Ces états viseraient à offrir une représentation du monde en même temps qu'à nous guider dans celui-ci en nous motivant. Millikan désigne les états en question par l'expression « pushmi-pullyu representations » (PPRs).

Son article de 1995 évoque les PPRs comme une possibilité, mais offre en fait assez peu d'arguments pour nous convaincre que les PPRs existent vraiment. Son premier argument consiste à dire que la sélection naturelle travaille souvent à partir de structures, mécanismes ou systèmes *moins* différenciés pour parvenir à d'autres structures, mécanismes ou systèmes *plus* différenciés. Selon cette logique, les états mentaux de désirs et de croyances auraient dû surgir à partir d'un état primitif ayant, dans une certaine mesure au moins, les caractéristiques de ces deux états à la fois. Sans doute est-ce vrai que la sélection naturelle procède souvent du moins différencié vers le plus différencié. Mais ce n'est pas toujours le cas et on ne peut pas exclure que les bases physiologiques qui sous-tendent les états de désirs et de croyance soient apparues de manière relativement indépendante. Seule l'histoire évolutive de ces bases physiologiques permettrait de trancher la question. Millikan n'offre rien en ce sens pour appuyer l'hypothèse des PPRs.

Millikan fait aussi remarquer que puisqu'un état d'un type est parfaitement inutile sans un état de l'autre type (un désir ne dirige pas sans croyance pour indiquer le moyen en vue de la fin posée ; et une croyance décrit « pour rien » si aucun désir ne vient s'y adjoindre), l'usage d'états purement descriptifs ou purement directifs implique la possession de capacités inférentielles pratiques relativement développées.

---

64. Millikan, 1995, p. 186.

L'usage de PPRs en revanche n'implique rien de tel. Une telle remarque suggère sans doute que les PPRs doivent exister de par le fait qu'un grand nombre d'animaux ne possèdent pas les capacités nécessaires à l'usage d'états « purs ». À cela, il faut répondre deux choses. D'abord, que d'affirmer qu'il existe entre états directifs et descriptifs cette relation de codépendance ne va pas à l'encontre de la THM, qui dit que la présence conjointe d'un désir et d'une croyance est nécessaire pour qu'il y ait motivation. Seulement, selon le humien, toute la force motivationnelle est dans le désir et la croyance ne sert qu'à mettre en lumière pour lui la présence de son objet, les caractéristiques de ce dernier, les moyens de l'atteindre, etc. La codépendance des états en question est entièrement compatible avec la décision théorique d'attribuer la force motivationnelle au seul désir. Ensuite, il se pourrait bien que l'action authentique, c'est-à-dire intentionnelle, présuppose bel et bien des capacités inférentielles suffisantes pour « emmagasiner de l'information pour laquelle on n'a aucune utilité immédiate (pure description), et pour représenter des buts qu'on ne sait pas encore comment atteindre (pure direction) »<sup>65</sup>. C'est ce que le reste de la présente section devrait faire ressortir.

Un dernier point évoqué par Millikan retiendra davantage notre attention ici. Il s'agit de la découverte d'un groupe de cellules nerveuses aux propriétés intéressantes : les « neurones miroirs »<sup>66</sup>. (La question n'est pas réglée de savoir si ces neurones existent chez l'homme, mais ceux-ci ont été observés chez plusieurs espèces de mammifères, dont les singes. Les spécialistes pensent qu'il est assez probable que

---

65. Millikan, p. 192.

66. *Ibid.*, p. 191. Parmi les premières publications traitant des neurones miroirs : Rizzolatti et al., 1981 et 1988 ; Gentilucci et al., 1983 et 1988, ainsi que Bruce et Golberg, 1985.

nous les retrouvions aussi dans le cerveau humain<sup>67</sup>.) Les neurones miroirs ont la propriété d'émettre des influx nerveux dans deux situations bien distinctes : lors de l'exécution d'une action simple et lors de l'observation d'un autre individu en train d'effectuer la même action. On a constaté chez les singes, par exemple, que ces neurones vont émettre un influx quand un singe effectue une action spécifique telle la saisie précise d'un objet, mais aussi quand une action équivalente (la saisie précise d'un objet) est effectuée par un individu (homme ou singe) que le singe en question est en train d'observer. Ces neurones semblent donc avoir pour fonction *à la fois* de générer une séquence motrice *et* de représenter des objets et des événements ; fonctions qui, d'habitude, sont respectivement associées aux états mentaux de désir et de croyance<sup>68</sup>. Les neurones miroirs pourraient-ils alors constituer le support matériel de la réalisation d'un authentique état mental de double direction d'ajustement ?

Il peut sembler que l'existence dans le cerveau humain d'un groupe de cellules ayant de telles propriétés soit de nature à jeter un sérieux doute sur la nécessité de la THM en tant que théorie générale de la motivation. De nature en tout cas à mettre hors jeu l'argument de l'uniformité théorique. L'argument, souvenons-nous, était que la THM s'était montrée nécessaire pour expliquer certaines actions simples. La THM était alors primée en tant que théorie générale de la motivation sur la base du fait qu'une généralisation de cette théorie constituait une option plus parcimonieuse que l'option d'attribuer le pouvoir de motiver (ou de générer un état intrinsèquement motivationnel) à l'état mental de croyance. Mais s'il existe des cas avérés où la

---

67. Voir Marc Jeannerod, *à paraître*.

68. Le système peut catégoriser différentes actions, selon qu'il s'agit, par exemple, de saisir délicatement avec les doigts ou bien d'agripper vigoureusement un objet avec la main entière.

motivation surgit d'un état unique qui présente à la fois les caractéristiques de la croyance et du désir, alors il n'y a bien sûr aucune économie à réaliser en généralisant l'explication humienne. Et les neurones miroirs pourraient bien constituer le support d'un tel état.

Une autre découverte en neurophysiologie semble toutefois interdire que l'on identifie la fonction motrice des neurones miroirs à la fonction motivationnelle du désir. Selon des études récentes<sup>69</sup>, les mouvements (c'est volontairement que je ne parle pas d'« actions ») que peuvent déclencher d'eux seuls les neurones miroirs sont nécessairement *non-intentionnels*. On constate que la schizophrénie induisait parfois chez les sujets qui en sont victimes l'illusion que certaines parties de leur corps sont contrôlées par des forces extérieures (le phénomène est connu en anglais sous le nom de « delusion of alien control »). Ces sujets ont l'impression que les gestes qu'ils posent n'originent pas d'eux-mêmes ; que les mouvements effectués par leur corps « leur arrivent » comme des événements auxquels ils assistent impuissants. Or, nous l'avons vu, pour qu'il y ait action à proprement parler, il faut qu'il y ait intentionnalité (et non simplement que des comportements « arrivent » d'eux-mêmes).

Les études montrent que les mouvements générés par les neurones miroirs sont normalement inhibés par un mécanisme neural situé dans le cortex préfrontal. Pour qu'il y ait action au sens propre du terme, c'est-à-dire comportement que l'agent revendique comme sien parce qu'intentionnel, il faut en effet la présence d'un « signal désinhibant [...] *qui représente la décision d'exécuter l'action* »<sup>70</sup>. La « décision »

69. Notamment Arbib, 2007.

70. *Ibid.*, p. 77. Certains lecteurs seront peut-être tentés d'attribuer à cette « décision » un statut épiphénoménal. Des expériences d'imagerie cérébrales montrent en effet que les neurones moteurs impliqués dans l'exécution d'un mouvement donné s'activent *avant même que l'individu n'ait*

supprime l'influence du mécanisme inhibiteur (« inhibe l'inhibition ») qui, chez un sujet normal, bloque habituellement les signaux moteurs qui sont générés de manière non-intentionnelle. Cela inclut les signaux envoyés par les neurones miroirs suite à l'observation d'actions exécutées par d'autres sujets. Ce qui caractérise les sujets victimes de schizophrénie dont nous venons de parler, c'est justement un déficit au niveau de la capacité d'inhiber les signaux générés de façon non-intentionnelle. Un cerveau sain inhibe ces signaux en permanence, sauf aux occasions où l'agent prend une « décision ». À ce moment, le cerveau sait que les signaux moteurs sont, pour ainsi dire, « commandés », et désactive le mécanisme inhibiteur.

Ainsi, les neurones miroirs pourraient peut-être constituer le support d'un état mental de croyance. Ils semblent en tout cas appropriés pour cela dans la mesure où leur patron d'activation sélectif indique qu'ils sont dotés d'une fonction représentative. Mais le potentiel moteur de ces neurones ne donne lieu à une action que si la « décision » de poser l'action en question *précède* l'émission de leurs influx. Les neurones miroirs sont tributaires de cette décision. Ils n'ont donc pas la capacité de motiver par eux-mêmes et, pour cela, ne peuvent pas constituer le support d'un état mental de désir (ou, plus généralement, d'un état mental de direction monde-esprit).

Évidemment, là où le neurophysiologue fait appel à une « décision », le humain trouvera l'occasion d'introduire la notion de « désir » ou d'« attitude conative ». Ainsi, contrairement à ce qu'aurait pu souhaiter l'anti-humain, la découverte des neurones

---

*conscience de prendre la décision d'effectuer ce mouvement* (voir à ce sujet Libet, 1991, pp. 27-28). Mais il me semble qu'il n'est pas nécessaire d'identifier la conscience qu'a un agent d'initier un mouvement à la « décision » dont parle Arbib. N'oublions pas non plus, pour ce qui suit, que le humain n'est pas contraint d'endosser une conception du désir qui ferait nécessairement de cet état mental un état conscient (comme c'est le cas pour la conception phénoménologique du désir par exemple).

miroirs et de leurs fonctions ne montre pas qu'il est nécessaire de postuler l'existence d'un état mental de double direction d'ajustement (ni même de son support matériel) ; c'est pourtant ce qui doit être montré si l'on veut contrecarrer l'argument humien de l'uniformité théorique<sup>71</sup>.

### 3.4 Conclusion du chapitre

Nous avons débuté notre examen des arguments à l'encontre de la THM par l'argument selon lequel la distinction supposée exister au niveau de la direction d'ajustement des désirs et des croyances est caractérisée de manière circulaire et non-informative. Le fait que les interprétations de l'idée de direction d'ajustement offertes par Smith et Zangwill réfèrent de manière essentielle à une « perception que non-p » constitue sans doute un vice majeur. Donner à « perception » le sens d'*apparence perceptuelle*, cela fait en sorte que la théorie donne parfois de mauvais résultats (en particulier dans les cas où l'agent suspend son jugement relativement à la véracité de la perception ou juge que celle-ci est trompeuse). Donner à « perception » le sens de « croyance », cela rend alors l'énoncé caractérisant l'asymétrie entre désirs et croyances complètement non-informatif. C'est la critique mise de l'avant par

---

71. Notons qu'un défenseur de la thèse des états mentaux de double direction d'ajustement qui voudrait s'aider de la découverte des neurones miroirs rencontrerait sans doute bien d'autres difficultés. La suivante n'est pas des moindres : certains neurones miroirs ne s'activent pas lors l'exécution de l'action qui les avait pourtant activés en tant qu'objet d'observation (ces neurones s'activent plutôt lors de l'exécution d'une ou plusieurs autres actions non reliées aux premières). Voir Fogassi & Gallese, 2002.

Humberstone et reprise par Copp et Sobel, qui insistent davantage sur le caractère circulaire des analyses de la notion de croyance offertes par Smith et Zangwill.

Notre attention s'est ensuite portée sur l'argument des croyances irrationnelles. Copp et Sobel font remarquer que celles-ci sont intuitivement considérées comme d'authentiques croyances, et ce malgré le fait qu'elles ne manifestent aucune tendance à disparaître à l'occurrence de perceptions de contenu contraire. Le dispositionalisme de Smith n'est donc pas approprié pour donner un sens plus précis à l'idée métaphorique de direction d'ajustement (lui qui ne considère pas comme une croyance un état mental qui, intuitivement, mériterait d'être considéré comme tel). Pour ce qui est de l'interprétation de Zangwill, nous avons vu que le problème venait de ce qu'on a envie de dire que la différence normative entre désirs et croyances découle d'une différence *non*-normative au sein de la nature de ces états mentaux ; et que les croyances irrationnelles, en particulier, illustrent le fait qu'une affirmation du type « la croyance est l'état dans lequel se trouve l'agent qui tient un contenu propositionnel donné pour vrai » (ou une autre définition essentiellement descriptive) constitue une caractérisation suffisante de l'état mental de croyance.

L'argument des états mentaux de double direction d'ajustement fut ensuite notre objet d'examen. Selon Pettit et quelques autres auteurs, le fait qu'il soit impossible à la croyance de contenu *p* de présenter simultanément les deux directions d'ajustement (cela implique en effet de manifester des dispositions incompatibles) ne montre pas qu'un état de double direction d'ajustement ne peut pas exister. C'est que, selon eux, on peut supposer qu'un tel état possède deux contenus distincts et que chaque contenu fasse l'objet d'une direction d'ajustement différente. Un tel « croisir » présenterait la

direction d'ajustement de la croyance face au contenu « x est bon » et la direction d'ajustement du désir face au contenu « x ». Il n'y aurait alors rien de problématique dans le fait que ce croisir ne tende pas à disparaître lorsque survient une perception dont le contenu est « non-x » (c'est plutôt à la perception qu'« il est faux que x est bon » que devrait réagir cet état mental).

Outre le fait que la thèse du croisir semble pour le moins ad hoc, l'existence d'un tel état mental reste hypothétique. Libre alors au humien de faire valoir qu'il est préférable de préserver l'uniformité de notre théorie (générale) de la motivation en n'admettant pas le croisir parmi les états mentaux intrinsèquement motivationnels. Millikan nous a toutefois indiqué qu'il existait peut-être une bonne raison de rompre l'uniformité en question : la découverte de neurones, les « neurones miroirs », qui semblent idéalement constitués pour servir de support à un authentique état de double direction d'ajustement. Ceux-ci paraissent effectivement avoir à la fois pour fonction de représenter les faits du monde et de « faire se mouvoir » l'agent. Et la fonction motrice serait (dans certains cas au moins) activée par la fonction représentative. On aurait donc la base matérielle pour quelque chose comme une croyance capable de motiver par elle seule.

Mais nous avons vu qu'en fait, « faire se mouvoir », c'est précisément ce que ces neurones avaient pour fonction d'accomplir. Les signaux moteurs générés par ces neurones ne donnent en effet lieu à des comportements intentionnels que s'ils ne sont eux-mêmes préalablement suscités par une « décision » de l'agent ; décision qui a pour effet de lever le mécanisme inhibiteur qui empêche normalement les signaux non-intentionnellement générés de provoquer une séquence comportementale. Les

neurones miroirs ne peuvent donc pas constituer le support d'un état intrinsèquement motivationnel et, par conséquent, ne peuvent pas non plus constituer le support d'un état de double direction d'ajustement. Ainsi l'anti-humien n'a pas à sa disposition ce qui semblait être une bonne raison pour nous faire accepter de rompre l'uniformité théorique de notre conception générale de la motivation.

## Conclusion Générale

À la lumière de l'analyse des différents arguments qui ont été ici invoqués en faveur ou à l'encontre de la THM, que peut-on dire de la situation de celle-ci ? Les arguments en sa faveur suffisent-ils à établir que nous avons besoin d'une théorie telle que la théorie humienne de la motivation ? Les objections soulevées contre elle suffisent-elles à montrer que la THM est fautive ou alors qu'elle souffre de quelque défaut majeur (incohérence, incomplétude, etc.) ?

Nous avons débuté en examinant ce en quoi consiste la THM et avons vu comment cette théorie était au moins superficiellement plausible, dans la mesure où elle se montrait capable d'offrir une explication pour la très vaste majorité des actions, même celles qui semblaient à première vue n'impliquer la présence d'aucun désir. On pense notamment aux actions commandées par les exigences de la morale lorsque celles-ci vont à l'encontre de nos inclinations premières. Cet examen de la plausibilité

superficielle de la THM nous a mené à considérer ce qui est tenu par certains (notamment Russ Shafer-Landau) comme une objection à part entière contre elle, à savoir le témoignage de la phénoménologie de la motivation morale. Un agent rapporte parfois ne pas avoir perçu la présence d'un désir parmi ses états mentaux au moment de poser une action prescrite du point de vue moral ; uniquement celle de quelque croyance à l'effet que cette action est obligatoire, ou juste, etc. De plus, il y a aussi des cas où un agent s'astreint à poser une action *non*-morale pour laquelle il semble n'avoir aucune inclination particulière. L'explication anti-humienne de tels cas, selon Shafer-Landau, est la plus « naturelle » et la plus « directe ».

Mais, prenant note des propos de Conie Rosati sur la phénoménologie de la motivation, nous avons vu qu'il était possible pour le humien d'expliquer des cas de cette espèce de manière satisfaisante. Le humien peut notamment invoquer un désir sous-jacent de l'agent, un désir plus fondamental dont la satisfaction constituerait sa véritable fin et aurait pour moyen l'action envisagée. Les soi-disant cas difficiles révélés par la phénoménologie de la motivation morale peuvent donc recevoir une explication humienne toute aussi directe et simple que celle fournie par l'anti-humien.

Nous avons ensuite abordé l'examen des arguments en faveur de la THM à proprement parler. L'argument des croyances identiques s'est avéré incapable d'offrir un soutien adéquat à la conclusion qu'il est censé appuyer, à savoir qu'il est faux qu'une croyance puisse motiver par elle-même (ou générer un état intrinsèquement motivationnel). La raison de ceci tient au fait que la thèse cognitiviste (anti-humienne donc) est suffisamment souple pour accommoder les cas où deux individus ont une même croyance au sujet de la valeur d'une chose, mais sont différemment motivés à

l'égard de celle-ci. Il est effectivement possible de soutenir que l'effet (sur la motivation) de l'introduction d'une croyance unique dans l'économie mentale d'un sujet dépend de l'*ensemble* des états mentaux (croyances et de désirs) parmi lesquels cette croyance est introduite. Deux sujets peuvent donc être différemment motivés à l'égard d'une chose, tout en ayant la même croyance quant à la valeur de la chose en question. Ainsi l'anti-humien s'engage-t-il seulement envers l'idée que les croyances motivent *ceteris paribus*, c'est-à-dire qu'elles motivent à condition que leur pouvoir motivationnel ne soit pas supplanté (*overridden*) par le pouvoir d'autres états mentaux présents dans l'économie mental de l'agent.

De surcroît, il n'y pas non plus de raison pour que deux agents partageant l'*ensemble* de leurs croyances soient identiquement motivés. C'est du moins l'idée que défend Margareth Little en affirmant qu'un même ensemble de croyances peut avoir différents effets sur la motivation d'un agent, selon que ce dernier subsume ou non ses croyances sous un « cadre conceptuel plus large ». Bien que ce cadre ne soit pas lui-même une croyance ou un ensemble de croyances, le processus de subsomption des croyances reste un processus entièrement cognitif - semblable à un *gestalt switch*. L'anti-humien dispose donc de réponses plausibles à l'argument des croyances identiques.

Nous nous sommes alors tourné vers la notion de direction d'ajustement, qui semblait pouvoir offrir la base d'un puissant argument en faveur de la THM. La notion pointe vers certaines caractéristiques prétendument essentielles des états de désir et de croyance, caractéristiques peut-être susceptibles de fonder la distinction entre ces deux types d'états mentaux et d'expliquer pourquoi seul un désir a la capacité de

motiver : si le contenu d'une croyance ne s'accorde pas avec les faits du monde, alors cela constitue une faute pour la croyance et celle-ci devrait s'« ajuster » pour qu'il y ait accord ; tandis que l'absence d'accord entre le contenu d'un désir et les faits du monde ne constitue pas une faute pour le désir qui devrait persister et disposer l'agent à « ajuster » les faits du monde de manière à obtenir l'accord en question. Désirs et croyances auraient donc des directions d'ajustement opposées et seul un désir aurait la direction appropriée pour rendre compte de ce que c'est qu'être motivé.

Mais le véritable défi consiste à donner un sens plus précis à cette idée métaphorique de direction d'ajustement. Smith en donne une interprétation dispositionaliste : alors que la croyance *que p* tend à disparaître à l'occurrence d'une perception *que non-p*, le désir *que p* tend à persister à l'occurrence de cette même perception – et dispose l'agent à faire en sorte que *p* se réalise. La présence d'un désir s'avère ainsi nécessaire pour motiver un agent à l'action puisqu'une croyance qui aurait le pouvoir de motiver devrait faire quelque chose d'impossible, soit manifester simultanément la tendance à persister à l'occurrence d'une perception de contenu contraire à son propre contenu (en tant qu'état moteur de l'action) et la tendance à disparaître à l'occurrence de cette même perception (en tant qu'il s'agit-là d'une caractéristique essentielle des croyances selon Smith). Pour sa part, Zangwill endosse une conception normative et soutient que la distinction entre désirs et des croyances vient de ce qu'une perception peut rationaliser une croyance sans pouvoir rationaliser un désir et qu'un désir peut rationaliser une intention (ou une action) alors que ce n'est pas le cas d'une croyance. La raison pour laquelle une croyance ne peut pas motiver est la même que chez Smith : un état cognitif tel une croyance devrait manifester

simultanément des dispositions incompatibles s'il avait la capacité de motiver. Seulement, Zangwill fait *découler* les dispositions manifestées par ces états mentaux de leur essence normative (alors que Smith en fait des caractéristiques *essentielles*).

À ce point de l'examen, j'ai affirmé que, bien que plusieurs auteurs semblaient croire que l'argument de la direction d'ajustement était le meilleur argument en faveur de la THM, j'étais pour ma part d'avis que l'argument<sup>72</sup> devait succomber à quelques objections simples (sur lesquelles nous allons nous attarder au chapitre suivant) et que, en l'état actuel du débat, la meilleure défense de la THM consistait plutôt dans l'argument de l'uniformité théorique<sup>73</sup>. Nous avons donc continué notre examen des arguments en faveur de la THM avec ce dernier argument.

Celui-ci consistait à dire que, puisque l'explication humienne est celle qui prévaut dans le domaine des actions non-morales, il est préférable (parce que plus parcimonieux) de généraliser cette explication aux actions morales que d'attribuer une nouvelle propriété à l'état de croyance – celle de motiver par elle-même ou de générer par elle-même un état mental intrinsèquement motivationnel. Cet argument présupposait deux choses en fait : d'une part, qu'il n'existe aucune action morale pour laquelle on ne dispose pas d'une explication humienne au moins aussi satisfaisante que l'explication anti-humienne. Cela semble être le cas : il n'y a, à ma connaissance, aucun cas d'action morale que le humien ne peut pas expliquer de manière plausible à

---

72. Ou plutôt : les *interprétations* de l'argument, puisque ces objections (l'objection de la circularité et de la non-informativité de la caractérisation de l'asymétrie entre désirs et croyances et l'objection des croyances irrationnelles) s'adressent spécifiquement aux interprétations en question.

73. Je tiens à préciser ici que la notion de direction d'ajustement me semble tout de même capturer quelque chose d'essentiel au sujet des croyances et des désirs. J'affirme seulement que les tentatives faites pour donner un sens non-métaphorique à cette notion se sont jusqu'ici avérées infructueuses. Ainsi, peut-être que d'autres interprétations seront avancées qui se montreront plus efficaces et ne succomberont pas aux objections soulevées ici.

l'aide de sa théorie de la motivation. Cela nécessite parfois qu'il fasse appel à des désirs que d'aucuns jugeront tarabiscotés. Mais les explications fondées sur ceux-ci sont somme toute assez vraisemblables. D'autre part, l'argument présuppose qu'il est vrai que la THM prévaut pour expliquer les actions non-morales ; ce que certains auteurs refusent d'admettre. Certains auteurs prétendent effectivement que les actions non-morales pourraient être motivées *sub specie boni*, c'est-à-dire que l'agent qui les pose se les représente (ou leurs conséquences) comme étant « bonnes » sous un rapport ou sous un autre.

Mais nous avons vu qu'il arrive qu'un agent soit motivé à accomplir une action non-morale sans avoir de croyance évaluative particulière : dans le cas d'actions posées par de jeunes enfants ou des animaux par exemple, ou dans le cas d'actions simples liées à nos dispositions physiologiques et psychologiques les plus fondamentales, telles se nourrir, avoir des rapports sexuels, éviter la douleur corporelle, etc. Ces actions *nécessitent* une explication humienne puisque l'état mental qui préside à leur accomplissement – et qui est donc intrinsèquement motivationnel – est dépourvu de la qualité représentative de la croyance (ou à tout le moins dépourvu de la capacité de représenter la valeur d'une action ou de ses conséquences).

L'argument de l'uniformité théorique semblait donc pouvoir fonctionner et offrir un appui d'une certaine force à la THM. Il restait toutefois à examiner les principales objections adressées à la THM, dont une en particulier (la possibilité d'une confirmation expérimentale de l'existence d'états mentaux de double direction d'ajustement) pouvait mettre en doute la valeur de l'argument de l'uniformité théorique – et, par conséquent, mettre en doute ce que nous savions à ce moment être l'unique

défense de la THM<sup>74</sup>, puisque l'argument de croyances identiques avait déjà échoué et que j'avais déjà annoncé que l'argument de la direction d'ajustement devait succomber aux objections qui lui seraient adressées.

Passons plus rapidement sur les arguments à l'encontre de la THM, étant donné que nous venons de clore le chapitre qui leur était consacré. L'argument du caractère circulaire et non-informatif des analyses de l'asymétrie entre désirs et croyances (en termes de direction d'ajustement) mettait en lumière le fait que les interprétations de Smith et de Zangwill caractérisent l'asymétrie en question en référant de manière essentielle à la direction d'ajustement de la croyance. Smith et Zangwill distinguent croyance *que p* et désir *que p* sur la base des relations (causales ou normatives) entretenues par ces états avec la « perception que non-p »<sup>75</sup>. On peut sans doute donner à « la perception que non-p » une interprétation « subdoxastique » qui lui conférerait le sens d'« apparence perceptuelle de contenu non-p ». Mais, dans ce cas, une partie des perceptions sera nécessairement laissée de côté puisque, on le sait bien, certaines perceptions semblent ne pas pouvoir provoquer la disparition des croyances dont le contenu est contraire au leur. Or en interprétant « la perception que non-p » de manière « doxastique » (i.e. en conférant à cette perception la direction d'ajustement de la croyance), les analyses de l'asymétrie deviennent pleinement circulaires (c'était la critique de Copp et Sobel). Et cette circularité ôte tout caractère informatif aux formules censées capturer la distinction entre désirs et croyances en termes de direction d'ajustement (c'était la critique d'Humberstone).

---

74. *Unique défense...* parmi les arguments ici considérés, qui ne sont pas les seuls comme je l'ai dit Je reviendrai à la toute fin de cette conclusion sur les raisons pour lesquelles j'ai examiné ces arguments plutôt que d'autres.

75. Ou, chez Zangwill : l'« expérience perceptuelle que p » (p. 195).

Mais un autre argument encore, l'argument des croyances irrationnelles, devait être soulevé face aux conceptions dispositionalistes et normatives de Smith et Zangwill. S'il est vrai que certains états mentaux peuvent être considérés comme des croyances alors qu'ils ne répondent pas aux perceptions de contenu contraire à leur propre contenu, alors le dispositionalisme ne révèle rien de *nécessaire* au sujet des croyances. Or les croyances irrationnelles (nous avons donné les exemples du *wishful thinking* et du délire paranoïaque), malgré leur absence de tendance à disparaître à l'occasion d'une perception de contenu contraire, sont considérées comme d'authentiques croyances. Le dispositionalisme de Smith est donc inapte à préciser le sens de l'idée d'une distinction des désirs et des croyances à partir de leur direction d'ajustement. En ce qui concerne le « fonctionnalisme normatif » de Zangwill, il n'est certes pas directement mis en doute par l'existence de croyances irrationnelles (dans la mesure en tout cas où il attribue correctement le statut de croyance aux états mentaux de ce type). Seulement, on pourrait être tenté de nier le statut fondamental des normes de rationalisation (celles-ci définissant la nature de chaque état mental selon Zangwill) et poser que celles-ci doivent découler de quelques propriétés non-normatives constitutives de l'essence de chaque état mental. L'existence des croyances irrationnelles confirmerait cette idée ; elles qui sont peut-être les plus à la limite de ne pas être considérées comme d'authentiques croyances et qui semblent quand même pouvoir être capturées par une définition strictement non-normative. Copp et Sobel par exemple, suggèrent qu'une croyance (rationnelle ou non) pourrait se définir comme l'état mental dans lequel se trouve un agent qui tient un contenu propositionnel donné pour vrai.

Nous avons ensuite fait comme si ces objections à l'argument de la direction d'ajustement n'étaient pas décisives et avons examiné l'objection des états mentaux de *double direction d'ajustement*. Qu'un état de croyance ne puisse pas assumer un rôle motivationnel sous peine de devoir présenter des directions d'ajustement opposées (et donc des dispositions incompatibles), cela est sans doute vrai dans le cas où l'état en question n'a qu'un seul contenu. Mais on peut poser que l'état motivationnel possède *deux contenus distincts*. Ce « croisir », qui serait un état à part entière et non la seule agglomération de deux états préexistants, présenterait un contenu pour chaque direction d'ajustement. Par exemple, « x » pourrait être le contenu sur lequel porte la direction d'ajustement monde-esprit (celle du désir), et « x est bon » le contenu sur lequel porte la direction esprit-monde (celle de la croyance). Ainsi, à l'occasion de la perception de contenu « non-x », il n'est pas nécessaire que l'aspect du croisir qui s'apparente à la croyance (et qui a pour contenu « x est bon ») tende à disparaître.

Sans doute que de postuler l'existence de croisirs peut paraître une manœuvre ad hoc mais, en tant que réelle possibilité théorique, cela permet à l'anti-humien de montrer que l'argument de la direction d'ajustement ne parvient pas en principe à exclure toutes les formes de croyances de l'ensemble des états mentaux motivationnels. Or, tant qu'il ne s'est pas avéré *nécessaire* d'introduire l'état de croisir dans notre typologie des états mentaux, le humien pourra faire valoir les considérations relatives à l'uniformité de la théorie de la motivation et tenir le croisir à l'écart de nos explications du phénomène de la motivation (en tant qu'il constitue une *possibilité* théorique seulement).

L'anti-humien pouvait à première vue espérer démontrer cette nécessité en s'aidant

de la découverte des « neurones miroirs ». Ces neurones semblent avoir pour fonction *à la fois* de représenter des objets ou des événements *et* de générer une séquence motrice ; fonctions habituellement associées aux états mentaux de désir et de croyance respectivement. Et c'est la fonction représentative qui, pour ainsi dire, « active » la fonction motrice : quand un singe se représente l'action accomplie par un autre singe, les mêmes neurones qui servent à la représentation envoient un influx moteur pour initier la même séquence comportementale. Les neurones miroirs paraissent donc pouvoir constituer le support matériel de la réalisation d'un authentique état mental de double direction d'ajustement.

Nous avons toutefois réalisé que ce mode de production de comportements ne pouvait pas convenir à l'explication d'actions intentionnelles. En effet, un mécanisme d'inhibition empêche normalement que les représentations (issues de la perception) ne soient immédiatement traduites en mouvements. Les actions à proprement parler, c'est-à-dire les comportements intentionnels, impliquent que l'agent lève d'abord ce mécanisme inhibiteur – par une « décision » disent les neurophysiologistes ; par un « désir » dira le humien. Les comportements directement issus de la représentation (c'est-à-dire sans que l'agent n'ait eu à lever l'inhibition qui empêchait que la représentation ne soit traduite en action) ne sont pas reconnus par le sujet comme lui appartenant. Cette nécessité qu'il y a, pour l'agent qui pose un acte intentionnel, de désamorcer le mécanisme inhibiteur indique l'échec de la représentation perceptuelle (et, dans la mesure où leurs fonctions – représentatives – s'apparentent, l'échec de la croyance) à produire seule la motivation et donc la possibilité d'une place pour le désir dans le schéma explicatif de l'action.

L'argument de l'uniformité théorique demeure donc une raison valable de préférer la THM à ses rivales anti-humiennes dans le débat sur la motivation. Mais, s'il s'agit d'une raison de favoriser la THM, l'argument de l'uniformité théorique n'est pas pour autant une raison *décisive*, au sens où il ne constitue pas une raison qui réglerait la question une fois pour toutes. L'avantage qu'il offre dans le débat n'est pas un acquis définitif : de nouveaux arguments pourraient toujours surgir et changer la donne. Seulement, dans l'état actuel du débat, l'argument semble suffire à faire pencher la balance du côté humien.

Comme nous l'avons dit, on pourrait accepter de rompre l'uniformité théorique pour des raisons *extérieures* au débat sur la motivation. Ces raisons impliquent des prises de positions substantielles – et généralement controversées – dans différents domaines de la philosophie. C'est pourquoi nous les avons tenues à l'écart. Mais sans doute serait-il intéressant de présupposer momentanément la vérité de ces positions théoriques afin de constater l'effet de ces raisons sur notre conception de la motivation.

L'une des raisons qui nous vient en premier à l'esprit est que la théorie hybride ainsi élaborée s'accorderait mieux avec la position réaliste en ontologie morale. Celle-ci gagne effectivement en plausibilité si l'on admet l'existence de croyances (morales) susceptibles de motiver par elles-mêmes : si l'on pense qu'il existe une réalité intrinsèquement morale et que la connaissance que nous avons de celle-ci peut nous motiver par elle-même (et non seulement par le truchement d'un désir comme le prétend le humien), il est alors nécessaire que nos croyances au sujet de cette réalité soient dotées du pouvoir de motiver – autrement le réaliste doit relier la réalité morale

à la motivation en faisant intervenir l'existence d'un désir « externe », tel celui de se conformer aux prescriptions de la morale. Évidemment, l'inverse serait de présupposer la vérité de l'anti-réalisme. Auquel cas – et en laissant de côté les « théories de l'erreur » (Mackie, 1977) – nos jugements moraux seraient l'expression de nos *seuls* désirs. Les soi-disant croyances dotées du pouvoir de motiver n'auraient alors aucun rôle à jouer ; donc la THM devrait être vraie.

Certaines prises de positions (encore là substantielles et controversées) relatives à la normativité et la rationalité pratique pourraient peut-être elles aussi constituer une base servant à mettre en doute la nécessité d'en appeler aux désirs pour expliquer les actions morales (voire les actions non-morales). C'est ce que font des auteurs tels Thomas Scanlon (Scanlon : 1998) et Jonathan Dancy (Dancy : 2000) me semble-t-il. Mais notons que, en ces matières, des prises de positions contraires pourraient aussi bien mener à d'autres arguments *en faveur* de la THM. J'ai en tête l'argument dit « de l'internalisme des raisons » qui, s'il pouvait légitimement être mis de l'avant, conférerait un appui considérable à la THM<sup>76</sup>.

Bref, si l'examen des arguments « internes » au débat sur la motivation nous indique que les considérations relatives à l'uniformité théorique constituent une raison de préférer la THM à ses rivales dans l'état actuel du débat ; on ne sait pas ce que l'examen des arguments qui s'appuient sur des thèses « externes » plus controversées pourrait nous indiquer au sujet de la bonne théorie de la motivation. Il pourrait en sortir de nouvelles raisons en faveur de la THM qui rendraient pour ainsi dire inutile de recourir à l'argument de l'uniformité théorique (en montrant qu'une croyance dotée

76. Cet argument est présenté dans Shafer-Landau, 2003, p. 128 ; mais il s'inspire en fait de réflexions de Bernard Williams, 1981, p. 102.

du pouvoir de motiver est une impossibilité conceptuelle par exemple). Mais il pourrait aussi en sortir des arguments à l'encontre de la THM qui montreraient que la rupture de l'uniformité est un prix qu'il est nécessaire (ou simplement préférable) de payer.

## Bibliographie

Altham, J. E. J., 1986, « The Legacy of Emotivism », dans G. Macdonald et C. Wright (dir.), *Fact, Science and Morality: Essays on A. J. Ayer's Language, Truth and Logic* », Oxford, Basil Blackwell, pp. 275-88.

Anscombe, G. E. M., 1957, *Intentions*, Oxford, Basil Blackwell.

Arbib, M.A., 2007 (juin), « Other faces in the mirror: a perspective on schizophrenia », *World Psychiatry*, 6 (2), pp. 75-78.

Blackburn, Simon, 1998, *Ruling Passions*, Oxford, Oxford University Press.

Bruce, C.J. et M.E. Golberg, 1985, « Primate frontal eye field. I. Single neurons discharging before saccades », *J. Neurophysiol.*, 53, pp. 603-635.

Copp, David et David Sobel, 2001, « Against Direction of Fit Accounts of Belief and Desire », *Analysis*, vol. 61, no. 1, pp. 44-53.

Copp, David et David Sobel, 2002, « Desires, motives, and reasons: Scanlon's rationalistic moral psychology », *Social Theory and Practice*, vol. 28, no. 2.

Dancy, Jonathan, 2000, *Practical Reality*, Oxford, Oxford University Press.

Davidson, Donald, 1980, *Essays on Actions and Events*, Oxford, Oxford University Press.

Fadigau, L., L. Fogassi, V. Gallese et G. Rizzolatti, 2000, « Visuomotor Neurons: Ambiguity of the Discharge or 'Motor' Perception? », *International Journal of Psychophysiology*, vol. 35, pp. 165-177.

Gentilucci, M., C. Scandolara, I.N. Pigarev, et G. Rizzolatti, 1983, « Visual responses in the postarcuate cortex area 6 of the monkey that are independent of eye position. » *Exp. Brain Res.* 50, pp. 464-468.

Gentilucci, M., L. Fogassi, G. Luppino, M. Matelli, R. Camarda, et G. Rizzolatti, 1988, « Functional organization of inferior area 6 in the macaque monkey: I. Somatotopy and the control of proximal movements », *Exp. Brain Res.*, 71, pp. 475-490.

Goldstick, Daniel, 2001, « Beliefs, Desires and Moral Realism », *Philosophy*, no. 81, pp. 153-160.

Hubel, David et Torsten Wiesel, 2005, *Brain and visual perception: the story of a 25-year collaboration*, Oxford, Oxford University Press.

Humberstone, L., 1992, « Direction of Fit », *Mind*, vol. 101, no. 401, pp. 59-83.

Hume, David, 1739-40, *Traité de la Nature Humaine*, Paris, Aubier, trad. A. Leroy (*A Treatise of Human Nature*, L. A. Selby-Biggy et P. H. Nidditch (dir.), Oxford, Clarendon Press, 1973), 1983.

Jeannerod, Marc, *à paraître*, « Consciousness of action and self-consciousness. A cognitive neuroscience approach », *à paraître* dans J. Roessler et N. Eilan (dir.), *Agency and self awareness : Issues in philosophy and psychology*, Oxford, Oxford University Press.

Kripke, Saul, 1982, *Wittgenstein on Rules and Private Language*, Harvard University Press.

Lenman, James, 1996, « Belief, Desire and Motivation: An Essay in Quasi-Hydraulics' », *American Philosophical Quarterly*, 33, 3, (juillet), pp. 291-301.

Libet, Benjamin, 1991, « Conscious vs. Neural Time », *Nature*, 352, 6330, pp. 27-28.

Little, Margaret O., 1997, « Virtue as Knowledge: Objections from the Philosophy of Mind », *Noûs*, no. 31, pp. 59-79.

Mackie, John, 1977, *Ethics: Inventing Right and Wrong*, Harmondsworth, Penguin.

McDowell, John, 1981, « Non-Cognitivism and Rule-Following », dans S. Holtzman et C. Leich (dir.), *Wittgenstein: To Follow a Rule*, London, Routledge, pp. 141-62.

Miller, Alexander, 2003, *An Introduction to Contemporary Metaethics*, Cambridge,

Polity Press.

Millikan, Ruth G., 1995, « Pushmi-Pullyu Representations », *Philosophical Perspectives*, vol. 9, AI, Connectionism and Philosophical Psychology, pp. 185-200.

Pettit, Philip, 1987, « Humeans, Anti-Humeans, and Motivation », *Mind*, vol. 96, no. 384, pp. 530-533.

Platts, Mark, 1979, *Ways of Meaning*, London, Routledge and Keagan Paul.

Price, Huw, 1989, « Defending Desire as Belief », *Mind*, vol. 98, no. 389, pp. 119-127.

Rizzolatti, G., Scandolara, C., Matelli, M., Gentilucci, M., 1981 « Afferent properties of periarculate neurons in macaque monkey. II. Visual responses », *Behav. Brain Res.* 2, pp. 147-163.

Rizzolatti, G., Gentilucci, M., 1988. « Motor and visual-motor functions of the premotor cortex. » In: Rakic, P., Singer, W. Eds., *Neurobiology of Neocortex*, Wyley, Chichester, pp. 269-284.

Rosati, Connie S., 2006, « Moral Motivation », *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/moral-motivation/>

Scanlon, Thomas, 1998, *What We Owe to Each Other*, Cambridge Mass., Harvard University Press.

Schueler, G. F., 1991, « Pro-Attitudes and Direction of Fit », *Mind*, vol. 100, no. 2, pp. 277-281.

Schueler, G. F., 2003, *Reasons and Purposes*, Oxford, Oxford University Press.

Shafer-Landau, Russ, 2003, *Moral Realism: A Defence*, Oxford, Clarendon Press.

Shafer-Landau, Russ, 2007, *Foundations of Ethics*, Oxford, Blackwell.

Smith, Michael, 1987, « The Humean Theory of Motivation », *Mind*, vol. 96, no. 381, pp. 36-61.

Smith, Michael, 1994, *The Moral Problem*, Oxford, Basil Blackwell.

Smith Michael, 2004, *Ethics and the A Priori: Selected Essays in Moral Psychology and Meta-Ethics*, New-York, Cambridge University Press.

Stampe, Denis, 1987, « The Authority of Desire », *Philosophical Review*, 96, pp. 335-381.

Stocker, Micheal, 1979, « Desiring the Bad: An Essay in Moral Psychology », *Journal of Philosophy*, pp. 738-753.

Tenenbaum, Sergio, 2006, « Direction of Fit and motivational Cognitivism », dans R. Shafer-Landau (dir.), *Oxford Studies in MetaEthics*, Oxford, Clarendon Press, pp. 235-264.

Van Roojen, Mark, 1995, « Humean Motivation and Humean Reason », *Philosophical Studies*, vol. 79, pp. 37-57.

Virvidakis, S., 1996, *La Robustesse du Bien. Essai sur le réalisme moral*, Nîmes, Jacqueline Chambon.

Williams, B.O., 1981, *Moral Luck*, Cambridge, Cambridge University Press.

Zangwill, Nick, 1998, « Direction of Fit and Normative Functionalism », *Philosophical Studies*, 91, pp. 173-203.