

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ANALYSE ET DÉFENSE DE L'APPROCHE ÉVALUATIVE EN THÉORIE DE
L'ACTION

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN PHILOSOPHIE

PAR
GABRIEL GÉLINAS

DÉCEMBRE 2017

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.10-2015). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

ANALYSE ET DÉFENSE DE L'APPROCHE ÉVALUATIVE EN THÉORIE DE L'ACTION

TABLE DES MATIÈRES

| | |
|--------------------------------------------------------|----|
| RÉSUMÉ | i |
| INTRODUCTION | 1 |
| CHAPITRE I | |
| L'APPROCHE ÉVALUATIVE | 10 |
| 1.1 Une présentation générale de l'approche évaluative | 10 |
| 1.2 La théorie « sans faille » de Tenenbaum | 13 |
| 1.3 La version modérée de l'approche évaluative | 20 |
| CHAPITRE II | |
| OBJECTIONS À L'APPROCHE ÉVALUATIVE | 30 |
| 2.1 L'argument des contre-exemples | 30 |
| 2.2 L'argument de la sophistication intellectuelle | 43 |
| 2.3 Le problème de la croyance | 51 |
| 2.4 Les raisons non normatives | 57 |
| CHAPITRE III | |
| VERSION OPTIMALE DE L'APPROCHE ÉVALUATIVE | 65 |
| 3.1 Pour une théorie sans faille | 65 |
| 3.2 Contre les valeurs anomiques | 74 |
| 3.3 Contre une théorie non séparatiste | 76 |
| CONCLUSION | 80 |
| BIBLIOGRAPHIE | 84 |

RÉSUMÉ

Ce mémoire porte sur l'approche évaluative en théorie de l'action. Cette théorie, aussi connue sous le nom anglais « *Guise of the Good Theory* », affirme que toute action intentionnelle est psychologiquement motivée par une appréhension de valeur de la part de l'agent. Le but de ce mémoire est d'analyser et de défendre une version « forte » de cette théorie. Dans le premier chapitre, nous exposerons les différentes versions de l'approche évaluative; la version radicale de Sergio Tenenbaum et les versions plus modérées de Joseph Raz et Katja Maria Vogt. Dans le second chapitre, nous examinerons les diverses objections faites à l'endroit de l'approche évaluative; l'argument des contre-exemples, l'argument de la sophistication intellectuelle, le problème de la croyance et l'argument des raisons non normatives. Nous verrons aussi comment les partisans de l'approche évaluative peuvent répondre à ces objections. Dans le troisième chapitre, nous défendrons une version sans faille de l'approche évaluative qui est également séparatiste, c'est-à-dire qui admet que certains états évaluatifs peuvent échouer à motiver à agir.

Mots clés : approche évaluative, « *Guise of the Good Theory* », action intentionnelle, valeur, motivation, évaluation, désir, raison pratique, psychologie morale.

INTRODUCTION

Dans le cadre de cette recherche, nous étudierons principalement le rôle que jouent les jugements de valeur et la perception du bien dans la motivation humaine. Lorsque nous sommes motivés à agir d'une certaine manière, est-ce toujours parce que nous voyons du bien dans ce que nous faisons? Y a-t-il toujours une appréhension de valeur sous-jacente à la motivation? Peut-on expliquer le comportement humain en partant du principe que les humains agissent uniquement lorsqu'ils perçoivent une valeur à ce qu'ils font? Le bien est-il la finalité de l'action humaine, ce vers quoi chacune de nos actions tend? Si oui, des exceptions à ce principe sont-elles possibles ou même concevables? Peut-on imaginer une action accomplie volontairement sans que l'agent ait en même temps une certaine compréhension que son action vaut la peine d'être faite, qu'elle est la bonne chose à faire? Dans ce mémoire, nous nous demanderons si la motivation humaine est intrinsèquement liée à ce que nous percevons comme étant le bien.

La découverte de principes fondamentaux derrière la motivation humaine permettrait à la philosophie de l'action d'acquérir une connaissance plus certaine de son objet. Si la motivation humaine était gouvernée par un principe tel que la poursuite du bien, alors il serait théoriquement possible de comprendre, d'expliquer et de prédire le comportement humain avec une grande certitude. En acquérant la connaissance de ce qui semble bon aux yeux d'une personne, nous serions du même coup en mesure de prédire comment elle va agir. Le processus de décision humain, qui semble aux premiers abords mystérieux et indéchiffrable, deviendrait une variable pouvant être appréhendée objectivement. De la même manière qu'en astronomie, la connaissance des lois de la physique nous permet de déduire de nouvelles informations à partir de données existantes, il serait possible de déduire, jusqu'à un certain point, les choix d'un être humain en fonction de ce que cette personne perçoit et ressent.

Ces aspirations sont peut-être irréalisables. Peut-être n'arriverons-nous jamais à comprendre pleinement ce qui motive l'action humaine. Une science de l'action est peut-être impossible à développer et à mettre en pratique, ou même une contradiction dans les

termes. Toutefois, si cette recherche nous permet d'en savoir plus sur la motivation humaine, alors elle est utile. Même si nous découvrons qu'il n'y a pas de principes guidant la volonté humaine et que les intentions demeureront toujours quelque peu mystérieuses, nous aurons quand même appris quelque chose d'important.

Plusieurs théories prétendent pouvoir expliquer pourquoi une personne agit. On peut les diviser en deux grandes catégories : les théories mentalistes et les théories non mentalistes. Les théories mentalistes donnent une explication psychologique de l'action. Elles cherchent à expliquer ce qui motive l'agent à accomplir une action en faisant référence à ses états mentaux. Les théories non mentalistes fournissent pour leur part une explication objective de l'action : certains faits non mentaux causent le comportement. Un exemple de théorie non mentaliste pourrait être l'idée selon laquelle tout ce qui est du domaine mental est causé par des phénomènes physiques et que notre conscience, nos sentiments et nos intentions sont tous déterminés par l'interaction de particules physiques à l'intérieur de notre cerveau. Sans que nous en ayons conscience, notre comportement serait déterminé par des causes physiques. Une autre théorie non mentaliste pourrait être que la génétique d'une personne, son environnement et son éducation déterminent les choix qu'elle fait et les actions qu'elle accomplit.

Ces théories, qu'elles soient vraies ou fausses, n'affectent pas notre recherche, puisque nous cherchons à savoir ce qui explique l'action humaine d'un point de vue subjectif. L'explication que nous recherchons doit faire référence au point de vue de l'agent. Les décisions d'agir sont des phénomènes mentaux et nous cherchons à savoir ce qui motive les décisions à l'intérieur même du cadre mental. Bien qu'on ne puisse douter que les facteurs mentionnés par les théories non mentalistes influencent le comportement d'une personne, nous ne saurions nous satisfaire d'une telle réponse. Une théorie qui s'appuie sur des faits physiques, biologiques ou sociologiques ne fournit pas le type d'explication que nous recherchons. Même s'il s'avère que le comportement des êtres humains est causé par le mouvement de particules physiques dans le cerveau, et conditionné par les gènes, l'environnement et l'éducation, une telle explication ne nous aide pas à connaître ce qui motive une personne à agir de son propre point de vue. Nous sommes à la

recherche d'une explication strictement psychologique qui ne se fonde sur rien d'extérieur à la conscience.

Pour satisfaire à nos exigences épistémologiques, l'explication que nous cherchons devra donc relever des théories mentalistes de l'action. Le débat contemporain met principalement en opposition deux théories mentalistes de l'action: la théorie causale et l'approche évaluative.¹ Dans ses grandes lignes, la théorie causale affirme que les actions intentionnelles sont directement causées par les états mentaux antérieurs, sans que ceux-ci doivent nécessairement inclure une perception ou un jugement de valeur. En d'autres termes, dans l'explication des actions, il est possible de fournir une réponse adéquate au pourquoi de l'action sans faire référence au bien ou à la valeur perçue par l'agent.² Les états mentaux X d'un agent auraient causé une action Y. Par exemple, l'action de l'agent de boire de l'eau s'explique par le fait que l'agent avait un désir de boire de l'eau et qu'il croyait que l'action de boire de l'eau satisferait ce désir. Il existe plusieurs variantes de la théorie causale, tout comme il existe plusieurs variantes de l'approche évaluative. Il est important de souligner qu'il n'est pas nécessaire pour une théorie causale de rejeter tout aspect évaluatif, et que certaines versions de la théorie causale peuvent admettre la perception du bien comme faisant nécessairement partie des états mentaux causant l'action. Toutefois, le cadre conceptuel de la théorie causale est tel qu'ajouter un élément évaluatif à la théorie causale a toutes les chances d'être vu comme une restriction superflue.³ C'est pourquoi, dans le cadre de cette recherche, nous désignerons comme théories causales uniquement les versions de cette théorie qui expliquent l'action sans faire référence à un jugement de valeur et nous considérerons l'approche évaluative comme totalement distincte de la théorie causale. Kieran Setiya défend une version de la théorie causale et nous étudierons cette position plus en détail dans le second chapitre.⁴

La théorie causale fait face à un problème majeur si elle tente de décrire le processus de décision en laissant de côté tout aspect évaluatif. Il est difficile de comprendre pourquoi

¹ Boyle et Lavin, 2010, p. 4.

² Setiya, 2010, p. 90.

³ Boyle et Lavin, 2011, p. 11.

⁴ Setiya, 2007, pp. 56-59.

une personne a accompli une action sans faire référence à une certaine compréhension de sa part que son action était la bonne chose à faire. Le rôle de la perception du bien dans notre processus de décision rationnelle est souvent négligé, mais à bien y réfléchir, on se rend compte qu'il est indispensable. Lorsque je dis, par exemple, que j'ai bu un verre d'eau parce que j'avais soif, la théorie causale présente le fait que j'avais soif comme la cause de ma décision de boire un verre d'eau. Cependant, une telle explication laisse implicite un élément nécessaire à la compréhension de l'acte de boire un verre d'eau en tant qu'acte rationnel. Il est en effet nécessaire d'inclure un élément évaluatif dans l'équation pour comprendre qu'il s'agissait bel et bien d'un acte rationnel. « J'ai soif, donc je bois de l'eau » n'est pas, en soi, un raisonnement. « J'ai soif, et étancher ma soif en buvant de l'eau est une bonne chose, donc je bois de l'eau », par contre, est un raisonnement intelligible. La présence d'un élément évaluatif dans le raisonnement, dans ce cas-ci la compréhension de la part de l'agent qu'étancher sa soif est une bonne chose, motive l'agent à vouloir accomplir son but. À la base de tout processus rationnel se trouve une certaine perception du bien et un désir de poursuivre ce bien. Cet élément demeure si souvent implicite dans nos raisonnements qu'il est facile d'oublier le rôle crucial qu'il joue ou même de nier son existence. On peut toutefois se convaincre de sa pertinence en examinant les cas où il est absent. Les actions qui semblent n'avoir aucune valeur sont justement celles que nous qualifions d'irrationnelles. Nous ne comprenons pas les motivations d'une personne qui agirait sans avoir en tête un quelconque but auquel on peut attribuer une certaine valeur. De telles actions nous semblent inintelligibles et irrationnelles.

Si cette critique de la théorie causale est valide, alors une théorie de l'action acceptable devra inclure un élément évaluatif dans son explication de l'action. La théorie qui fera l'objet de cette recherche remplit cette exigence. Il s'agit de l'approche évaluative, aussi connue sous le nom de « Guise of the Good Theory ». Plusieurs variantes de l'approche évaluative existent, mais elles ont toutes en commun d'affirmer que la perception du bien joue un rôle dans le processus de décision qui mène à l'action. La perception de valeur a le potentiel d'être le « pont » reliant la réalité objective et la réalité subjective en ce qui concerne la motivation humaine. Si nous acceptons l'approche évaluative, alors nous

devons accepter que l'explication la plus directe de l'action intentionnelle fasse référence à quelque chose de subjectif, soit la perception de valeur ou du bien de la part de l'agent. Cependant, il doit être possible de connaître objectivement quelles choses sont en mesure de causer une perception de valeur chez un être humain sous quelles conditions. Il serait peut-être plus facile d'expliquer le comportement indirectement, par le biais de ce qui cause une perception de valeur chez l'humain, plutôt qu'en tentant d'expliquer directement les actions comme étant causées par l'environnement, l'éducation ou encore le mouvement des particules physiques dans le cerveau. La perception de la valeur ou du bien a le potentiel d'être un point où la conscience et le monde physique se rejoignent lorsqu'il est question d'intention et de volonté. Tout comme il est possible de dire que l'interaction d'ondes lumineuses avec l'appareil oculaire cause la perception de la couleur sans que celle-ci soit réductible à des propriétés physiques, il serait possible de dire que certains stimuli causent une perception subjective du bien, ce qui motive l'action, sans que la motivation elle-même soit réductible à des propriétés physiques.

L'approche évaluative est prometteuse, mais plusieurs questions doivent encore être élucidées avant de pouvoir formuler une théorie cohérente. Par exemple, il faut déterminer ce à quoi le concept de bien réfère. G.E. Moore affirme que nous savons tous intuitivement ce qu'est le bien, que nous avons tous la même idée en tête lorsque nous parlons du bien, mais que celui-ci est impossible à définir en des termes autres que lui-même.⁵ D'autres philosophes modernes, ne se laissant pas décourager, ont proposé plusieurs définitions du bien. Pour certains, ce qui est bien est ce qui est plaisant. Pour d'autres, il s'agit de ce qui est désiré.⁶ Nous ne prétendons pas prendre position sur la question du bien dans cette recherche. Cependant, pour les besoins du présent travail, nous pourrions nous reposer sur l'idée de Moore selon laquelle nous avons tous une idée relativement similaire en tête quand nous pensons au bien. Une chose est certaine, cependant. Notre conception du bien, dans le cadre de la présente recherche, ne signifie pas ce que nous concevons généralement comme le bien moral. Si c'était le cas, l'approche évaluative devrait nier que des gens puissent commettre des actions

⁵ Moore, 1903, pp. 45-46, 58.

⁶ Sumner, 1996, ch. 3-5.

immorales, ce qui est absurde. Prenons, par exemple, le cas d'un voleur. Il peut sembler bon au voleur de dérober un objet précieux même si cela constitue un acte immoral. Le bien, dans ce cas, est perçu comme étant quelque chose d'avantageux pour l'agent lui-même, et non pas comme ce qui est moral ou bon pour le plus grand nombre. C'est aussi un acte qui semblera à la fois bon du point de vue du voleur et mauvais du point de vue de la victime. Il est donc possible pour une certaine action d'apparaître à la fois bonne et mauvaise en fonction du point de vue que l'on adopte.

Il est important de souligner également que cette recherche ne s'intéresse pas tant à la nature du bien qu'à la perception qu'en ont les êtres humains. Lorsqu'on dit que le bien motive quelqu'un à agir, il s'agit d'une explication psychologique. Que le bien existe de manière objective ou pas, les êtres humains ont néanmoins une conception de ce qui est bon et valable et ils forment des jugements évaluatifs en se basant sur ces notions. C'est sur cette perception subjective du bien que se fonde l'approche évaluative. L'approche évaluative offre une explication subjective de la volonté et de l'action. Nous pensons que seule une explication subjective est en mesure d'offrir le genre de compréhension directe du pourquoi de l'action intentionnelle que nous recherchons, puisque sans cet aspect subjectif, nous ne serions pas en mesure de faire le lien entre la théorie et notre expérience, et ce lien est nécessaire lorsque nous souhaitons étudier quelque chose comme la motivation, qui est du domaine de la conscience.

Un autre problème que l'approche évaluative doit résoudre concerne la manière dont l'individu entre en contact avec le bien ou la valeur. En utilisant le terme « perception du bien », nous laissons entendre que la valeur et le bien se manifestent directement à l'esprit, qu'ils ne sont pas la conclusion d'un quelconque raisonnement ou le résultat d'un jugement délibéré. Pourtant, il n'est pas évident que cela soit toujours le cas. La question de savoir comment l'individu entre en contact avec la valeur, si la valeur est découverte par l'individu, ou créée par lui, est une question importante à laquelle l'approche évaluative devra répondre. Nous pourrions aussi nous demander quels types de choses peuvent avoir de la valeur. La valeur est-elle perçue dans l'action elle-même, dans les conséquences de l'action, dans l'état du monde à la suite de l'accomplissement de

l'action, ou est-elle contenue dans les états mentaux des individus qui sont rendus heureux ou malheureux par l'action? La valeur, bien que perçue subjectivement, peut-elle être quelque chose d'objectif, ou est-elle uniquement subjective? Nous ne répondrons pas à toutes ces questions en détail dans la présente recherche, mais il est utile de noter qu'elles ont une importance pour la théorie.

L'idée selon laquelle nos actions sont motivées par le bien est une idée ancienne. On la retrouve déjà énoncée par Platon dans le *Ménon*, lorsque Socrate affirme que personne, en réalité, ne désire le mal, et que tous désirent le bien. Ceux qui font le mal le font involontairement en croyant faire le bien, car personne ne veut se nuire à soi-même et être misérable.⁷ Aristote, dans l'*Éthique à Nicomaque*, affirme que tous les choix et toutes les actions ont pour but l'accomplissement d'un bien quelconque, et que le bien est ce vers quoi toutes les choses tendent.⁸ Cette idée est demeurée populaire à travers les siècles, spécialement au sein de la scolastique médiévale, si bien qu'aujourd'hui certains auteurs y font référence comme à la théorie scolastique,⁹ tandis que d'autres la désignent tout simplement comme l'approche classique ou traditionnelle.¹⁰ L'approche évaluative a toutefois fait l'objet de plusieurs attaques qui ont mis en doute sa validité dans les derniers siècles. On peut faire le lien entre le déclin de l'approche évaluative d'une part et l'émergence de la méthode scientifique et le désir de remettre en questions les idées anciennes, jugées trop simplistes, d'autre part. Dans un article intitulé « Goodness and Desire », Matthew Boyle et Douglas Lavin avancent l'idée que la raison principale derrière l'opposition contemporaine à l'approche évaluative est qu'elle semble rendue obsolète par la théorie causale de l'action, qui est aujourd'hui plus communément acceptée.¹¹ Cette théorie explique l'action simplement en faisant référence à l'état psychologique antérieur de l'agent, et ne nécessite pas de faire référence à ce qui semble bon à l'agent dans le raisonnement menant à l'action. En acceptant une telle explication de l'action, nous reléguons du même coup l'approche évaluative à l'état de bagage métaphysique superflu.

⁷ Platon, *Ménon*, 77b-78b.

⁸ Aristote, *Éthique à Nicomaque*, 1094a.

⁹ Tenenbaum, 2007, p. 21.

¹⁰ Raz, 1999, p.23; Setiya, 2007, p. 21.

¹¹ Boyle et Lavin, 2010, pp. 3-4.

Nous défendrons toutefois l'argument qu'une telle approche que la théorie causale de l'action n'est pas réellement intelligible si elle ne fait aucune référence à la valeur. Il y a toujours une place en théorie de l'action pour l'approche évaluative, si nous sommes en mesure de la justifier. Plusieurs philosophes défendent encore aujourd'hui cette théorie, qu'ils désignent le plus souvent par l'expression anglaise « *Guise of the Good Theory* ». Le débat autour de la question de la motivation et de l'action est encore bien vivant, et l'approche évaluative a ses détracteurs aussi bien que ses défenseurs. Le but de ce mémoire est de faire un survol du débat actuel et de tenter de dégager, au terme de nos efforts, une version de la théorie qui soit à la fois plausible, explicative, et qui résiste aux objections courantes contre l'approche évaluative.

Dans le premier chapitre, nous présenterons deux versions distinctes de l'approche évaluative. Nous résumerons chacune d'entre elles et identifierons leurs forces et leurs faiblesses. La première est une version « sans faille », proposée par Sergio Tenenbaum, qui affirme que motivation et évaluation concordent parfaitement dans tous les cas. Selon cette version radicale de la théorie, on ne peut pas désirer une chose sans que celle-ci nous apparaisse bonne d'une certaine manière, et vice versa. La seconde version, proposée par Joseph Raz et Katja Maria Vogt, est une version de l'approche évaluative qui admet des exceptions et qui est considérée comme une version plus modérée de la théorie. Dans la plupart des cas, nous sommes motivés par une appréhension du bien, mais cette version ne rejette pas la possibilité d'avoir des motivations perverses, donc de faire le mal volontairement, contrairement à la pensée socratique. Des cas d'acrasie lorsque l'on perçoit de la valeur dans une action sans être suffisamment motivé à l'accomplir, par paresse ou faiblesse de la volonté et des cas d'acédie lorsque l'on n'est pas du tout motivé par la perception de ce qui est bien sont admis par cette version de la théorie. Il y a aussi certaines actions qui sont trop triviales pour mériter un effort de délibération. Dans ces cas-là, les actions ne sont pas motivées par le bien.

Dans le second chapitre, nous présenterons les objections principales ayant été adressées à l'approche évaluative ainsi que la réponse des défenseurs de l'approche évaluative à

celles-ci. Nous présenterons également des arguments originaux visant à défendre l'approche évaluative contre chacune des objections. Nous commencerons par l'argument classique des contre-exemples avancé par Michael Stocker et invoqué par pratiquement tous ceux qui se sont penchés sur la question. Selon cet argument, nous ne pouvons pas être uniquement motivés par le bien, puisque certaines personnes commettent des actes mauvais ou cruels ou peuvent refuser de faire ce qui leur semble bon, par paresse ou pour d'autres raisons. Stocker affirme également que l'approche évaluative est trop simpliste pour pouvoir décrire la complexité de notre esprit. Nous présenterons ensuite un argument de David Velleman. Selon Velleman, nous ne pouvons pas lier le fait de désirer quelque chose au fait de concevoir cette chose comme bonne parce que certains agents, comme les animaux et les jeunes enfants, sont capables de désirer, mais sont incapables de former un jugement évaluatif du type « X est bien ». Il est donc possible d'avoir l'un sans l'autre, contrairement à ce que la théorie avance. Nous présenterons finalement deux arguments de Kieran Setiya contre l'approche évaluative. Le premier est qu'une théorie de l'action plausible doit pouvoir satisfaire au critère de la croyance, qui est d'expliquer comment l'agent croit qu'il fait X au moment où il fait X, ce que l'approche évaluative ne fait pas. Le second argument est qu'il est tout à fait possible pour un agent d'expliquer pourquoi il accomplit une action sans évoquer une quelconque évaluation de cette action comme bonne, ce qui indique que les raisons d'agir peuvent être purement explicatives et ne faire aucune référence au bien et à la valeur.

Dans le troisième chapitre, nous discuterons les détails d'une version optimale de l'approche évaluative. Dans un premier temps, nous soutiendrons que l'approche évaluative doit être sans faille, c'est-à-dire qu'elle ne doit pas admettre d'exceptions au principe que les actions intentionnelles sont motivées par une évaluation. Par la suite, nous rejeterons l'idée que des valeurs anomiques, comme le mal, puissent motiver à agir, la seule valeur capable de motiver étant le bien. Finalement, nous nous positionnerons en faveur d'une version séparatiste de l'approche évaluative qui affirme que la motivation à agir implique toujours la présence d'une évaluation, mais qu'il n'est pas vrai que toutes les évaluations motivent à agir.

CHAPITRE I

L'APPROCHE ÉVALUATIVE

1.1 Une présentation générale de l'approche évaluative

L'approche évaluative est, comme l'explique Sergio Tenenbaum, une théorie à propos de l'action intentionnelle et de la motivation humaine.¹² La thèse générale de l'approche évaluative, aussi connue sous le nom de « Guise of the Good Theory », stipule que tout ce qu'un agent désire et veut accomplir est compris par lui comme étant un bien. Plusieurs versions de l'approche évaluative existent, mais une version minimaliste de la théorie affirmerait que dans les cas classiques d'action intentionnelle, un agent accomplit une action X seulement s'il conçoit d'une certaine manière que X est bon. Toutes les versions de l'approche évaluative doivent au moins accepter cette idée fondamentale. Joseph Raz donne une définition de la théorie qui avance que puisque les actions intentionnelles sont des actions faites pour des raisons perçues par l'agent et que les raisons sont des faits qui établissent aux yeux de l'agent que l'action a de la valeur, alors les actions intentionnelles sont faites parce que l'agent a la croyance qu'elles ont de la valeur.¹³ L'approche évaluative postule donc une connexion essentielle entre motivation et évaluation à l'intérieur de l'esprit humain. Elle rationalise les actions intentionnelles en montrant comment elles paraissent bonnes du point de vue de l'agent.

Il est important de souligner que l'approche évaluative n'est pas une forme d'internalisme moral. L'internalisme moral est une théorie qui affirme que si un agent juge qu'une action est moralement bonne, alors il est intrinsèquement motivé à l'accomplir. La différence entre cette théorie et l'approche évaluative se situe au niveau de ce que nous entendons comme étant le bien. Comme nous l'avons expliqué précédemment, le bien que postule l'approche évaluative n'est pas ce qu'on appelle communément le bien moral. Pour éviter la confusion, le type de bien auquel nous faisons référence est parfois appelé le bien *simpliciter* ou encore le bien « tout court ». L'internalisme moral fait face à

¹² Tenenbaum, 2013, p. 1.

¹³ Raz, 2008, p. 1.

de sévères critiques puisque certains agents, comme les psychopathes, ne semblent pas posséder de sens moral et ne sont pas motivés à faire ce qui est moralement bon. Ces agents pourraient toutefois être motivés par ce qui leur semble bon *simpliciter*, comme le voleur qui dérobe un objet précieux par convoitise. Les actions des agents immoraux ne sont pas exclues de l'approche évaluative comme elles le sont de l'internalisme moral.

Nous ne voulons pas non plus faire de l'approche évaluative une théorie triviale qui stipule que le bien se définit tout simplement comme ce qui est désiré par un agent. Dans ce cas, l'approche évaluative serait nécessairement vraie puisque l'on pourrait dire que toutes les actions intentionnelles sont faites en vue du bien parce que le bien est ce que l'agent désire. Une telle théorie ne nous apprendrait toutefois rien de nouveau et n'aurait aucune valeur philosophique. Nous tentons de découvrir un fait réel concernant la motivation humaine qui nous permettrait de mieux expliquer les actions intentionnelles.

Nous voulons aussi éviter de construire une version trop forte de la théorie, qui serait facilement réfutable. Une telle théorie pourrait consister à dire que nous sommes seulement motivés à faire les choses que nous jugeons bonnes après une réflexion approfondie.¹⁴ Cette idée est manifestement fautive parce que nous ne prenons pas toujours le temps de réfléchir de cette manière avant d'agir et aussi parce que certains agents, comme les animaux, ont la capacité de désirer et d'agir sans avoir la capacité de former un jugement réfléchi de type « X est bon ». Tenenbaum affirme qu'une version plus raffinée de l'approche évaluative stipulerait que nous sommes motivés par une apparence, une sensation ou une perception du bien, plutôt que par une croyance formelle de type « X est bon ». Un désir est, selon Tenenbaum, un état mental dans lequel quelque chose nous apparaît comme un bien.

Une autre clarification opérée par Tenenbaum insiste sur le fait que l'approche évaluative n'affirme pas simplement que l'agent accomplit une action X parce qu'il pense que X en soi est une bonne chose, mais plutôt parce qu'il croit que le fait d'accomplir X fera

¹⁴ Tenenbaum, 2013, p. 5.

survenir le bien. Le bien est, selon Tenenbaum, le but constitutif de l'action.¹⁵ En accomplissant une action, l'agent a pour but de faire survenir un bien quelconque en conséquence de l'action. Par exemple, le voleur dérobe un objet de valeur non pas parce qu'il juge que l'action en soi est bonne, mais plutôt parce qu'il souhaite faire survenir un bien, qui est de s'enrichir. Le défi pour l'approche évaluative sera de démontrer que les actions intentionnelles partagent un but commun, le bien, sans toutefois faire de cette théorie une explication triviale qui définit *a priori* le bien comme étant le but de toute action. L'approche évaluative doit démontrer que le bien est quelque chose de substantiel qui joue un rôle nécessaire dans la motivation à agir. Si l'approche évaluative relève avec succès le défi et qu'elle est en mesure de survivre aux objections qui la menacent, alors, comme l'affirme Joseph Raz, elle a le potentiel d'être la clé de voûte permettant de faire le lien entre les théories de la valeur, de la normativité et de la rationalité, et notre compréhension de l'action intentionnelle.¹⁶

Dans le reste de ce chapitre, nous examinerons en détail deux versions distinctes de l'approche évaluative. La première, la théorie « sans faille » de Tenenbaum, peut être considérée comme la version radicale de la théorie. Les versions développées par Raz et Vogt sont plus modérées. Il existe deux différences principales entre ces deux versions. La première est que pour Tenenbaum, l'apparence de bien implique la motivation, et vice versa. Autrement dit, l'apparence de bien est suffisante pour causer une motivation, et toute motivation est causée par une apparence de bien. Pour Raz, c'est plutôt les jugements de valeur qui impliquent la motivation, mais l'inverse n'est pas nécessairement vrai. Lorsque nous percevons de la valeur dans quelque chose, nous sommes généralement motivés à agir en ce sens. Cependant, cela ne signifie pas que chacune de nos actions intentionnelles est orientée vers le bien. Contrairement à Tenenbaum, Raz pense que certaines actions intentionnelles peuvent être accomplies sans que l'agent ne voie une quelconque valeur en elles. La deuxième différence entre les deux théories se situe au niveau des exceptions. La théorie de Tenenbaum ne tolère pas d'exceptions au principe que les apparences de bien motivent. Pour Raz et Vogt, la situation est

¹⁵ *Ibid.*, p. 6.

¹⁶ Raz, 2008, p. 34.

différente. Même si la plupart du temps, les apparences de bien nous motivent à agir, il est parfois possible d'être motivé par des valeurs anomiques telles que le mal. Ces deux différences font en sorte que la théorie de Tenenbaum est la version la plus « radicale » de la théorie, tandis que les versions de Raz et Vogt sont perçues comme plus « modérées ».

1.2 La théorie « sans faille » de Sergio Tenenbaum

En développant sa théorie, Sergio Tenenbaum avait bien conscience des objections couramment dirigées à l'encontre de ce qu'il appelle la « théorie scolastique traditionnelle ». Le point de vue exprimé par Emmanuel Kant dans la *Critique de la raison pratique*, que nous désirons uniquement ce que nous concevons comme bien,¹⁷ est l'ancêtre de l'approche évaluative. Cependant, ce point de vue présente des ambiguïtés et est vulnérable à certaines objections lorsqu'il est soumis à une analyse plus approfondie. Tenenbaum pense toutefois que l'idée de Kant est en grande partie vraie et son but est donc de présenter une version « sans faille » de la théorie scolastique qui fasse concorder parfaitement les éléments de motivation et d'évaluation dans l'esprit humain et qui soit immunisée contre les critiques courantes.¹⁸

La théorie de Tenenbaum est non séparatiste et non subjectiviste.¹⁹ Il convient de définir ces termes. Une théorie est séparatiste si elle considère qu'il peut exister des états motivationnels qui ne sont pas évaluatifs, ou encore des états évaluatifs qui ne sont pas motivationnels. Une théorie non séparatiste comme celle de Tenenbaum affirme le contraire. Autrement dit, chaque fois que nous sommes motivés à faire quelque chose, c'est que nous évaluons que cette chose est bonne, et chaque fois que nous évaluons qu'une chose est bonne, nous sommes motivés à agir en ce sens. De plus, pour Tenenbaum, la force des états motivationnels doit être proportionnelle à la force des contenus évaluatifs. Autrement dit, nous ne pouvons pas être très fortement motivés à faire quelque chose qui nous apparaît seulement comme un bien mineur. Plus une chose

¹⁷ Tenenbaum, 2007, p. 1.

¹⁸ *Ibid.*, pp. 13-14.

¹⁹ *Ibid.*, pp. 2-3.

nous apparaît bonne, plus nous sommes motivés, et vice versa. Une théorie est non subjectiviste, selon Tenenbaum, si elle accepte que le bien est quelque chose de normatif et d'indépendant, nous permettant de juger objectivement le contenu de nos désirs. Par opposition, une position subjectiviste affirmerait que le bien n'est rien d'autre que ce que nous désirons, qu'il ne s'agit pas d'une notion indépendante qui puisse guider notre action. Au mieux, une position subjectiviste peut rendre l'approche évaluative trivialement vraie. Au pire, elle la rend tout simplement fausse.

La thèse de Tenenbaum peut être formulée de la manière suivante : désirer quelque chose est identique à concevoir cette chose comme bonne selon une certaine perspective évaluative.²⁰ Certaines des notions de cette théorie méritent une explication plus approfondie. Le désir est un état motivationnel où quelque chose apparaît bon à l'agent, mais ce n'est pas une intention. Le lien entre le désir et l'intention est que le désir est causé par une apparence de bien, tandis que l'intention consiste en un jugement décisif qu'une action X est la bonne chose à faire. L'intention mène directement à l'action, tandis que le désir est une apparence qui ne cause pas directement l'individu à agir. Le médiateur entre le désir et l'intention est, selon Tenenbaum, la raison pratique et une conception générale du bien.²¹

Le terme « concevoir » mérite aussi quelques clarifications. Tenenbaum utilise le concept de « concevoir quelque chose comme bon » pour illustrer le fait que le bien ne se présente pas tout à fait comme une connaissance, mais plutôt comme une apparence. Cette distinction est importante parce qu'elle reflète le caractère faillible de l'agent et laisse la porte ouverte à l'erreur. En effet, une chose peut nous apparaître bonne, mais cela ne garantit pas que cette chose le soit réellement. Les apparences peuvent être fausses ou illusoire. Par exemple, un gâteau au chocolat peut sembler délicieux et me donner envie d'en manger, mais cela ne signifie pas nécessairement qu'il est bon pour moi d'en manger. Malgré les apparences, il a peut-être mauvais goût, peut me rendre malade, etc. Le fait que le gâteau m'apparaisse bon ne constitue pas une connaissance

²⁰ *Ibid.*, p. 21.

²¹ *Ibid.*, p. 76.

que le gâteau est véritablement bon pour moi; c'est seulement une impression. Il est possible d'avoir des désirs contradictoires, comme le désir de manger du gâteau parce qu'il semble bon et le désir de ne pas en manger pour éviter de devenir obèse. Le plaisir de manger du gâteau et le fait de ne pas être obèse sont deux choses qui apparaissent bonnes, mais les actions qui mènent à ces deux résultats sont incompatibles. Dans ce cas, les désirs sont contradictoires et il est impossible de satisfaire les deux en même temps. Si les désirs étaient une connaissance du bien, et pas seulement une apparence, de telles contradictions seraient impossibles. Cependant, puisque les désirs sont des apparences, et que par conséquent ils peuvent être illusoire, la contradiction entre différents désirs est possible et même courante. Même si je juge, par mon raisonnement, que manger du gâteau est mauvais, le gâteau peut continuer de m'apparaître bon. Je peux donc continuer de le désirer même si mon jugement dicte que je ne dois pas le manger. Dans un tel cas, le désir de manger du gâteau est encore motivé par une apparence de bien, qui est l'aspect délicieux du gâteau et cette apparence de bien fait en sorte que je continue de le désirer même si ma raison me dicte que je ne dois pas en manger.

Tenenbaum fait une analogie entre le bien et la vérité, en disant que tout comme la raison théorique a pour but d'arriver à la vérité, la raison pratique a pour but de faire survenir le bien.²² Concevoir ou imaginer que X est vrai me fait percevoir X comme étant vrai d'une certaine manière, mais cela ne m'oblige pas à croire que X est vrai. De la même manière, concevoir une chose comme étant bonne ne m'oblige pas à croire que cette chose est réellement bonne. La raison pratique doit utiliser les croyances et les apparences qui sont à sa disposition pour réussir à former un jugement que X est bon, et donc former l'intention de faire X. Pour y parvenir, la raison pratique doit utiliser une conception générale du bien afin de discerner le bien réel de l'illusion. Tenenbaum définit l'idée de conception générale du bien comme une conception pratique et réfléchie de ce qui est bon, formée par un agent compétent à l'aide de l'expérience et de la raison.²³ Cette conception générale du bien permet à l'agent de délibérer sur ce qu'il doit ou ne doit pas faire, elle lui indique les désirs qu'il vaut la peine de poursuivre et ceux qu'il vaut mieux

²² *Ibid.*, p. 21.

²³ *Ibid.*, pp. 53-54.

ignorer. C'est en se fiant à sa conception générale du bien que l'agent peut, par exemple, s'abstenir de manger un gâteau au chocolat qui semble délicieux s'il juge que, toutes choses considérées, il vaut mieux maintenir une alimentation saine.

La théorie de Tenenbaum fait face à un autre défi, celui de déterminer quelles actions sont explicables par l'approche évaluative. N'existe-t-il pas des actions faites sans aucune raison, que l'on ne peut pas expliquer par le fait que l'agent percevait du bien dans ce qu'il faisait? On peut penser aux actions accidentelles, qui ne sont pas le fruit de l'intention. Par exemple, si je renverse accidentellement mon verre d'eau et que, par conséquent, l'eau se répand sur le sol, il serait manifestement faux de dire que le fait de renverser mon verre m'apparaissait comme un bien. Au contraire, je n'ai jamais voulu renverser mon verre. C'était un accident. Les actions involontaires comme celle-ci ne sont pas couvertes par l'approche évaluative. Mais existe-t-il une autre catégorie d'action, qui inclurait des actions faites volontairement, mais sans que l'agent ne perçoive une quelconque valeur dans leur accomplissement? Tenenbaum pense que oui, et il utilise comme exemple les actions pathologiques, celles que l'agent accomplit sans en avoir le désir ou même sans reconnaître qu'elles ont un but ou une fonction.²⁴ Imaginons une personne qui ne puisse s'empêcher de déplacer et de replacer une pile de livres. Lorsqu'on lui demande si elle voit de la valeur dans son action, elle répond qu'elle n'en voit aucune, mais qu'elle ne peut pas s'empêcher de déplacer les livres. Cette personne accomplit pathologiquement cette action sans en éprouver le désir.

Les actions pathologiques sont controversées, car il est difficile de savoir si, véritablement, aucun désir n'est présent dans le processus menant à agir. Le problème des actions pathologiques et d'autres types d'actions similaires est toutefois assez important pour inciter Tenenbaum à introduire une nouvelle notion dans sa théorie. Il s'agit du critère d'intelligibilité.²⁵ Pour qu'une action soit explicable par la théorie scolastique, elle doit être intelligible. Pour cela, il faut que l'action ait un certain sens aux yeux de l'agent et qu'elle aspire à une certaine rationalité. Pour savoir si une action répond au critère

²⁴ *Ibid.*, p. 32-33.

²⁵ *Ibid.*, p. 33. De l'anglais « intelligibility requirement ».

d'intelligibilité, nous pouvons utiliser une méthode proposée par Anscombe, qui consiste à se poser la question : « pourquoi? ». ²⁶ Le « pourquoi? » d'Anscombe est une notion importante utilisée par plusieurs auteurs s'intéressant à la théorie évaluative de l'action, tels que Tenenbaum et Setiya. Pour qu'une action satisfasse à ce critère, il faut que l'agent soit en mesure de dire pourquoi il a accompli l'action. Il doit être capable de fournir une explication quant au motif de son action. Par exemple, si j'ouvre une porte pour entrer dans une pièce et qu'on me demande pourquoi j'ai ouvert la porte, je peux répondre que c'est parce que je voulais entrer dans la pièce. Par contre, je ne peux pas expliquer pourquoi j'ai renversé le verre d'eau, parce que c'est une action accidentelle. L'explication fournie en réponse au « pourquoi? » d'Anscombe ne peut pas être purement causale, par exemple « J'ai renversé le verre d'eau parce que ma main l'a accidentellement accroché et que le mouvement a fait basculer le verre ». L'explication doit rendre l'action intelligible en tant qu'action volontaire; elle doit être une explication qui révèle la raison pour laquelle l'agent a choisi d'accomplir un geste. Une action à propos de laquelle une réponse satisfaisante au « pourquoi? » d'Anscombe peut être donnée est une action pouvant faire l'objet de l'approche évaluative, selon Tenenbaum. Une action pathologique comme celle décrite plus haut ne répond pas au critère d'intelligibilité et n'est donc pas explicable par la théorie.

La théorie scolastique doit aussi pouvoir expliquer un autre type d'action, à savoir l'action volontaire, mais non réfléchi. Imaginons, par exemple, que j'aie un verre d'eau à portée de main. Tout d'un coup, sans réfléchir, je décide d'en boire une gorgée. Ce n'est certainement pas une action accidentelle. Ce n'est pas non plus une action pathologique, parce que je peux répondre au « pourquoi? » d'Anscombe. En effet, j'ai décidé de boire une gorgée d'eau parce que j'en avais envie. Mais boire une gorgée d'eau était-elle la meilleure chose à faire à ce moment-ci? Était-il mieux pour moi de ne pas boire, ou encore, d'attendre cinq minutes avant de boire cette eau? Je ne sais pas; je n'ai pas réfléchi en profondeur à cette question. J'ai instinctivement décidé, en voyant le verre d'eau, d'en boire une gorgée. Il s'agit d'une action volontaire, mais non réfléchi.

²⁶ Anscombe, 1963, pp. 5-8; Setiya, 2010, p. 89; Tenenbaum, 2007, p. 33.

Comment cette action peut-elle être expliquée par la théorie scolastique si j'admets moi-même que je n'ai pas réfléchi aux différentes possibilités et que j'ai agi impulsivement?

Malgré le fait que je n'ai pas pris la peine de réfléchir, mon action de boire de l'eau demeure intelligible. Je peux tout à fait l'expliquer en invoquant le bien, qui est dans ce cas-ci le fait d'étancher ma soif. Le problème, c'est que je n'ai pas pris la peine de délibérer pour décider s'il s'agissait de la meilleure chose à faire. Tenenbaum résout ce problème en disant que dans certaines circonstances, notre conception générale du bien dicte qu'il vaut mieux ne pas prendre le temps de pleinement délibérer à propos de chaque action.²⁷ Il est préférable de ne pas perdre trop de temps à réfléchir à certaines actions triviales comme boire de l'eau et de simplement les accomplir quand nous en avons envie. Contrairement à Raz et Vogt, Tenenbaum est d'avis que même les actions triviales et irréfléchies sont explicables par l'approche évaluative.

L'approche évaluative peut aussi prendre en compte l'idée d'une conception déontologique du bien. Une théorie déontologique est une théorie morale qui affirme que certaines actions sont moralement interdites, peu importe les conséquences qu'elles peuvent générer. Par exemple, une approche déontologique pourrait affirmer que sacrifier des personnes innocentes est toujours une mauvaise chose. Or, on peut imaginer des scénarios où sacrifier la vie d'une personne peut sauver la vie de plusieurs autres personnes. Dans certains cas, on pourrait penser qu'une action déontologiquement interdite constitue la meilleure chose à faire dans les circonstances. Comment alors réconcilier une approche déontologique avec la théorie scolastique proposée par Tenenbaum? La réponse de celui-ci est de dire qu'il est non seulement possible, mais plausible, que les théories déontologiques soient elles-mêmes orientées vers le plus grand bien possible.²⁸ Adopter une théorie déontologique, c'est croire que le fait de suivre un certain nombre de règles morales fera subvenir le plus grand bien de manière générale, même si cela n'est pas le cas dans certaines situations particulières. Ainsi, même si un agent peut penser qu'une action moralement interdite pourrait faire subvenir plus de bien

²⁷ Tenenbaum, 2007, p. 95.

²⁸ *Ibid.*, pp. 199-200.

qu'une action moralement permisible dans une situation donnée, sa croyance que le fait d'obéir à un code déontologique en toutes circonstances fera subvenir le plus grand bien possible peut le convaincre de ne pas dévier de son code moral peu importe la situation dans laquelle il se trouve. L'agent déontologique agit donc en accord avec la théorie scolastique, puisqu'il juge que le fait de se conformer aux règles déontologiques fera survenir un plus grand bien que le bien immédiat pouvant découler d'une action moralement interdite.

Rien dans la théorie « scolastique » de Tenenbaum ne présuppose que les actions d'un agent doivent être réellement bonnes ou qu'elles ont nécessairement une valeur objective. Il est tout à fait possible pour un observateur neutre de désapprouver des actions d'un agent particulier, de penser qu'elles n'ont aucune valeur ou qu'elles sont objectivement mauvaises. S'il existe une valeur objective en fonction de laquelle il est possible de juger les actions particulières, certaines actions peuvent être objectivement mauvaises et être tout de même explicables par la théorie « scolastique ». De telles actions demeurent intelligibles si elles apparaissent bonnes à l'agent au moment où il les accomplit. L'apparence de bien est une impression subjective qui affecte l'agent, mais une telle apparence peut être fausse ou encore n'avoir aucune valeur de vérité si le bien objectif n'existe pas. La théorie « scolastique » reflète le caractère imparfait de l'agent, qui n'a pas nécessairement un accès épistémologique au bien objectif, mais qui agit selon ses impressions subjectives et sa raison imparfaite. Puisque la théorie « scolastique » admet des erreurs de la part de l'agent, elle admet aussi que l'agent puisse éprouver des regrets. Tenenbaum présente deux situations dans lesquelles l'agent peut éprouver des regrets.²⁹ Ainsi, un agent peut regretter avoir mangé un gâteau au chocolat qui semblait délicieux, si le gâteau s'avère être rassis et goûter mauvais. Il peut aussi regretter avoir mangé le gâteau s'il juge après coup qu'il eût été mieux de faire un effort pour respecter sa diète. Ainsi, d'une part, les résultats de l'action peuvent se révéler décevants comparativement au désir initial, et d'autre part, le désir lui-même peut être jugé comme ne valant pas la peine d'être satisfait. D'une manière comme de l'autre, il est possible pour l'agent de regretter une action qu'il a accomplie.

²⁹ *Ibid.*, p. 28.

La théorie « scolastique » de Tenenbaum présente plusieurs avantages. Tout en demeurant sans faille, c'est-à-dire en n'admettant pas d'exceptions au principe que l'apparence de bien motive à l'action, elle permet néanmoins d'expliquer la multiplicité des attitudes humaines, comme le regret, le fait d'opter entre deux biens incompatibles, d'opter pour le moindre mal, d'agir sans réfléchir, etc. Elle prend aussi en compte le caractère imparfait de l'agent, qui agit souvent sur la base de fausses croyances ou d'apparences illusoire. Tenant compte des objections les plus courantes contre la théorie « scolastique » traditionnelle, Tenenbaum a développé une version de la théorie plus à même de se défendre contre ces objections. Que le désir soit lié à une apparence de bien, plutôt qu'à un jugement réfléchi, rend la thèse « scolastique » beaucoup plus crédible. Le défi pour Tenenbaum consiste à montrer que le principe selon lequel l'apparence du bien nous motive à agir s'applique à toutes les actions volontaires et qu'il ne survient pas seulement de manière accidentelle. Nous nous demanderons dans le deuxième chapitre si la théorie sans faille de Tenenbaum survit aux objections à son endroit.

1.3 La version modérée de l'approche évaluative

Dans cette section, nous présenterons deux versions plus modérées de l'approche évaluative, développées par Joseph Raz et Katja Maria Vogt. Bien que ces deux théories soient distinctes, elles ont en commun un point important : elles tolèrent des exceptions au principe selon lequel les actions intentionnelles sont motivées par le bien. Contrairement à Tenenbaum, Raz et Vogt adoptent un point de vue séparatiste. Certaines actions volontaires ne sont pas motivées par une perception de valeur ou de bien. Pour ces deux auteurs, le fait de tolérer des exceptions à la règle rend l'approche évaluative plus plausible et lui permet de mieux rendre compte de certaines situations concrètes. Cela accommode, par exemple, les cas de dépression sévère où l'agent ne souhaite pas que sa vie aille bien ou encore les actes d'agression motivés par la colère où faire du mal est précisément ce qui est désiré.

Raz présente sa version de l'approche évaluative dans deux articles intitulés « The Guise of the Good » et « Agency, Reason and the Good ».³⁰ La thèse de Raz postule premièrement l'existence d'un lien conceptuel fort entre raison d'agir et valeur. Une raison se définit, selon lui, comme un fait établissant qu'une action a de la valeur.³¹ Pour être en mesure de motiver à agir, une raison doit être comprise par l'agent. Elle apparaît donc sous la forme d'une croyance. La croyance qu'il existe une raison d'accomplir une certaine action rend cette action désirable aux yeux de l'agent, elle fait paraître l'action comme valant la peine d'être accomplie. Elle confère donc de la valeur à l'action. Ainsi, lorsqu'un agent accomplit une action pour une raison, il est conceptuellement nécessaire qu'il croie que son action possède une certaine valeur. C'est à cause de cette valeur qu'il pense avoir une raison d'accomplir l'action, et c'est pourquoi il l'accomplit effectivement. Une action faite pour une raison sera toujours, sauf dans certains cas exceptionnels, faite *sub specie boni*, c'est-à-dire en vue du bien. Les actions faites pour des raisons sont donc explicables par l'approche évaluative.

Un autre aspect important de la thèse de Raz est le caractère implicite du lien qui existe entre raison et valeur. Il n'est pas nécessaire pour l'agent de comprendre explicitement que son action a de la valeur pour être motivé à l'accomplir.³² Au contraire, même les agents qui ne possèdent pas les concepts de bien et de valeur agissent parce qu'ils perçoivent de la valeur dans ce qu'ils font. Lorsqu'un agent agit pour une raison, il tient pour acquis le bien-fondé de son action. Il approuve l'action lorsqu'il l'accomplit, et cette approbation implique que l'agent pense que son action a de la valeur. Selon Raz, cette dimension implicite fait en sorte que l'approche évaluative est en quelque sorte auto confirmée par les faits.³³ Lorsque nous voyons quelqu'un accomplir une action intentionnellement et que cette personne a une raison d'accomplir cette action, alors nous devons nécessairement conclure que cette personne voyait un certain bien dans son action. Si elle n'avait pas cru que son action avait une certaine valeur, elle n'aurait pas eu raison de l'accomplir, et par conséquent elle ne l'aurait pas fait. Ceci s'explique par le

³⁰ Raz, 1999; Raz, 2008.

³¹ Raz, 2008, p. 1.

³² *Ibid.*, pp. 5-7.

³³ *Ibid.*, p. 7.

fait que le lien entre raison et valeur est un lien conceptuel présent dans la définition même des concepts de valeur et de raison et que ce lien existe, peu importe ce qu'une personne peut croire à ce sujet. En fait, si une personne accomplit un geste conscient et délibéré puis affirme que cette action a été faite sans raison ou encore qu'elle n'a aucune valeur, nous devons considérer que cette personne se trompe ou est confuse conceptuellement, tout comme quelqu'un qui croirait qu'une personne est célibataire, mais qui refuserait de croire que cette même personne est non mariée.³⁴ Lorsqu'un agent accomplit une action et que son action est motivée par une raison, alors nous devons naturellement considérer que son action a été faite *sub specie boni*.

Nous avons vu dans les sections précédentes que l'approche évaluative s'applique aux actions intentionnelles et que toute action faite intentionnellement est accomplie en vue d'une certaine valeur ou d'un certain bien. Pour Joseph Raz, la question est plus complexe. En effet, toutes les actions intentionnelles ne sont pas explicables par la théorie, puisque Raz tolère des exceptions à la règle. Avant d'examiner les exceptions possibles, nous allons voir quels sont les critères pour qu'une action appartienne à la catégorie d'actions explicables par la théorie. Raz introduit le concept d'intention indépendante. Une intention indépendante est une intention qu'il est possible de conserver même lorsque nous ne sommes pas en train d'accomplir l'action en question.³⁵ Par exemple, je peux avoir l'intention d'écrire un livre, et cette intention se maintient même lorsque je ne suis pas activement en train d'écrire ce livre. Même dans les moments où mon attention est portée sur autre chose et que je n'ai pas à l'esprit mon projet d'écrire un livre, je reconnais que j'ai tout de même cette intention et que celle-ci se maintient aussi longtemps que je n'ai pas terminé ou abandonné mon projet. Lorsqu'on a une intention indépendante de faire une action, cela implique que l'on a une raison d'accomplir cette action.³⁶ Ainsi, selon Raz, toute action qui découle d'une intention indépendante est couverte par l'approche évaluative.³⁷

³⁴ *Id.*

³⁵ *Ibid.*, p. 9.

³⁶ *Ibid.*, p. 16.

³⁷ *Id.*

Toutefois, les intentions indépendantes ne sont pas le type d'intention servant à produire directement l'action. Les intentions qui servent à générer directement les actions intentionnelles sont les intentions incorporées (*embedded intentions*).³⁸ Illustrons ce point à l'aide d'un deuxième exemple. Admettons que j'aie une intention indépendante de me rendre à l'arrêt d'autobus. Ceci implique de marcher une certaine distance sur le trottoir pour me rendre jusqu'à l'arrêt. L'acte de marcher est constitué par toute une série de mouvements, effectués sur une certaine période de temps, et ces mouvements nécessitent une certaine forme d'intentionnalité pour être accomplis. Je dois, d'une certaine manière, vouloir accomplir ces mouvements pour qu'ils s'accomplissent. Ces mouvements ne sont pas directement générés par autant d'intentions indépendantes, mais tombent tous sous la gouverne d'une seule intention indépendante, qui est de me rendre à l'arrêt d'autobus. Chacun de ces mouvements, par contre, dépend d'une intention incorporée qui permet de mettre le corps en mouvement afin d'accomplir l'action voulue. Ainsi, l'action qui consiste à marcher vers l'arrêt d'autobus découle d'une intention indépendante, mais n'est pas générée par une action indépendante. Elle résulte des intentions incorporées. Il est important de comprendre que lorsque Raz affirme que les actions accomplies avec une intention indépendante sont explicables par l'approche évaluative, il dit quelque chose à propos des intentions incorporées qui génèrent l'action et qui sont placées sous la gouverne d'intentions indépendantes, et non pas quelque chose à propos des intentions indépendantes elles-mêmes. Les intentions indépendantes sont certainement adoptées pour des raisons, et donc *sub specie boni*, mais elles ne causent pas directement l'action. Raz doit démontrer que les actions elles-mêmes, accomplies par le biais des intentions incorporées, sont accomplies parce que l'agent voit du bien en elles.

Les intentions incorporées qui découlent d'une intention indépendante peuvent être de deux types. Elles peuvent être des intentions dérivées ou non dérivées.³⁹ Les intentions incorporées non dérivées sont celles qui résultent naturellement et directement de l'intention indépendante qui les gouverne. Reprenons l'exemple de l'arrêt d'autobus mentionné précédemment. Dans cet exemple, l'acte de marcher sert directement à

³⁸ *Ibid.*, p. 10.

³⁹ *Ibid.*, pp. 10-11.

accomplir le but de l'intention indépendante, qui est de se rendre à l'arrêt. Cette action est une condition nécessaire à l'accomplissement de l'intention indépendante. Ainsi, si l'intention indépendante est motivée par une raison, et que cette intention requiert nécessairement une action donnée, alors l'intention incorporée qui permet d'accomplir cette action est motivée par cette même raison, et l'action est nécessairement accomplie parce que l'agent voit de la valeur en celle-ci. Les intentions incorporées non dérivées sont donc explicables par l'approche évaluative.

Les intentions incorporées dérivées posent un problème plus sérieux à la théorie. Ce sont les intentions qui ne sont pas directement supportées par une intention indépendante, mais qui surviennent en tant que conséquences collatérales d'actions faites avec une intention non dérivée. Imaginons que lorsque je marche vers l'arrêt d'autobus, je me retrouve face à un écureuil sur le trottoir. Je sais que si je continue d'avancer, je vais effrayer l'écureuil et que celui-ci va courir se réfugier dans l'arbre le plus proche. Je n'ai aucune raison de penser qu'effrayer les écureuils constitue une chose bonne en soi. Au contraire, je préfère généralement les laisser tranquilles. Cependant, il se trouve qu'un écureuil est présentement sur mon chemin et que si je continue d'avancer vers l'arrêt d'autobus, cela aura pour conséquence nécessaire de l'effrayer. J'ai une raison de marcher vers l'arrêt d'autobus, mais pas de raison particulière d'effrayer l'écureuil. Cependant, comme je n'ai pas l'intention de m'arrêter pour un écureuil, je continue d'avancer volontairement, tout en sachant que l'écureuil sera effrayé. J'effraie donc l'écureuil volontairement, même en n'ayant pas de raison indépendante de le faire, et même si je ne pense pas que cette action ait en soi une quelconque valeur. Ma raison d'effrayer l'écureuil est tout simplement que c'est une conséquence nécessaire d'une action faite avec une intention indépendante, qui est de marcher vers l'arrêt d'autobus. L'intention d'effrayer l'écureuil est une intention incorporée dérivée de l'intention incorporée non dérivée qui est de marcher vers l'arrêt d'autobus. Raz affirme qu'on ne peut pas considérer les intentions non dérivées et dérivées tout à fait séparément. Il faut les considérer comme un tout.⁴⁰ Ici, le fait d'effrayer l'écureuil ne suffisait pas à me convaincre de m'arrêter. En continuant d'avancer tout en sachant que cela allait avoir

⁴⁰ *Ibid.*, p. 13.

pour résultat d'effrayer l'écureuil, j'ai décidé que ma raison de me rendre à l'arrêt d'autobus était assez forte pour continuer malgré cette conséquence indésirable. L'intention incorporée dérivée n'a peut-être pas de valeur indépendante, mais lorsqu'elle est associée à l'intention non dérivée pour former un ensemble, alors même l'intention dérivée peut être expliquée par l'approche évaluative. Comme le dit Raz, les intentions incorporées non dérivées sont explicables par la théorie, et les intentions incorporées dérivées sont explicables par le fait qu'elles sont dérivées.⁴¹

Nous avons vu que dans la théorie de Raz, les actions explicables par l'approche évaluative sont les actions faites avec une intention indépendante. Ceci laisse la porte ouverte à un certain nombre de cas exceptionnels non explicables par la théorie. Raz distingue trois types d'exceptions à l'approche évaluative. Il y a, premièrement, les actions faites volontairement, mais sans raison. Raz classe dans cette catégorie les actes insignifiants auxquels on porte typiquement peu attention.⁴² Des actes anodins tels que se passer la main dans les cheveux ou taper des doigts sur la table sont certes des actes intentionnels, mais ils ne sont pas nécessairement faits avec une intention indépendante. Ces actes peuvent être accomplis sans raison véritable et peuvent donc constituer des exceptions authentiques à l'approche évaluative. Il est intéressant de noter que la frontière entre ce qui constitue un acte insignifiant et ce qui constitue un acte fait avec une intention indépendante n'est pas tout à fait définie par Raz. Celui-ci considère qu'une action telle que prendre une gorgée d'eau peut être accomplie sans une intention indépendante de la part de l'agent.⁴³ Cet exemple marque une différence concrète entre la thèse de Raz et celle de Tenenbaum, qui considère que l'acte de boire une gorgée d'eau, même lorsqu'il est irréfléchi, demeure tout de même intelligible et s'inscrit par conséquent dans la catégorie des actions faites en vue du bien.

La deuxième catégorie de cas exceptionnels inclut les actions qui semblent être faites pour une raison, mais dont l'intentionnalité est réduite. Ceci inclut principalement les

⁴¹ *Id.*

⁴² *Ibid.*, p. 15.

⁴³ *Ibid.*, p. 9.

actions expressives, c'est-à-dire les actions faites sous le coup de l'émotion.⁴⁴ Admettons que, sous le coup d'une intense colère, je frappe le mur devant moi de toutes mes forces. On pourrait être tenté de dire que je n'avais aucune raison d'agir ainsi, que frapper le mur n'est pas une action intelligible. Cependant, il est indéniable que j'ai frappé le mur volontairement. Et n'avais-je pas, sous l'emprise de la colère, la très nette impression de devoir l'exprimer en agissant de la sorte? N'avons-nous pas, dans de telles situations, une certaine impression que de laisser s'exprimer l'émotion est la bonne chose à faire? Un autre exemple, développé par Raz, illustre la quasi-rationalité qui anime les actions expressives. Imaginons qu'une mère, sous l'emprise de la colère, s'en prenne à son enfant et le gifle.⁴⁵ Même si on peut dire que la mère n'a pas de raison de frapper son enfant et qu'elle va certainement ressentir de la culpabilité à cause de ce geste, ce n'est pas un hasard qu'elle ait pris pour cible l'objet de sa colère, c'est-à-dire l'enfant. Elle était suffisamment en contrôle de ses moyens pour diriger l'expression de sa colère sur l'objet de celle-ci, tout comme celui qui frappe un mur peut parfois demeurer assez lucide pour ne pas choisir une surface particulièrement dure. En fait, les actions expressives peuvent presque être considérées comme des actions faites pour des raisons. L'agent approuve typiquement ces actions au moment où il les commet, même s'il peut les regretter par la suite, et il est souvent en mesure d'expliquer pourquoi il a agi de cette manière en invoquant l'émotion qui a motivé l'acte.⁴⁶ Cependant, malgré le fait que les actions expressives soient intentionnelles et démontrent un certain degré de rationalité, elles ne sont pas soumises à l'approche évaluative. Raz explique que ces actions démontrent une certaine perte de contrôle de la part de l'agent et que même si elles sont en grande partie intentionnelles, elles ne le sont pas entièrement. Typiquement, les actions expressives sont moins bien coordonnées que les actions normales. Une personne qui frappe le mur parce qu'elle est en colère risque de mal contrôler son mouvement et de se blesser en frappant trop fort. L'intentionnalité derrière ces actions n'est pas parfaite. Le pouvoir décisionnel de l'agent se trouve affaibli lorsqu'il est sous l'emprise de l'émotion, et donc les actions expressives sont des cas exceptionnels qui ne se conforment pas à l'approche

⁴⁴ *Ibid.*, pp. 16-19; Raz, 1999, pp. 36-44.

⁴⁵ Raz, 1999, p. 41.

⁴⁶ *Ibid.*, p. 43.

évaluative, mais qui ne la réfutent pas non plus.⁴⁷ Les actions pathologiques, telles que la cleptomanie et les actes accomplis sous hypnose, peuvent aussi être considérés comme des actions dont l'intentionnalité est réduite, et donc comme des exceptions à l'approche évaluative, selon Raz.

Il y a finalement les actions anomiques, c'est-à-dire les actions faites en vue du mal. Lorsque nous prenons le mal comme raison d'agir, nous agissons pour des raisons anomiques.⁴⁸ Raz considère que les actions anomiques sont des cas dégénérés, des inversions de la norme qui est d'agir pour des raisons qui établissent que l'action est bonne. Agir en vue du mal implique toujours agir en vue d'une certaine valeur, sauf que cette valeur est inversée. Il s'agit du mal plutôt que du bien. Ces inversions de la norme sont des exceptions qui ne réfutent pas l'approche évaluative, selon Raz, parce qu'il est toujours possible de dévier d'une norme établie et que même dans le cas des actions anomiques, la valeur demeure un élément central de l'action.⁴⁹ Raz considère qu'il est normal d'avoir autant de cas exceptionnels, puisque l'approche évaluative ne cherche pas à être sans exception. De plus, une théorie qui ne tolère pas d'exceptions est une théorie dont on doit se méfier parce qu'elle est probablement fautive. Plutôt que de minimiser l'importance des contre-exemples, il vaut mieux les examiner afin de déterminer pourquoi ils ne se conforment pas à la théorie, ce que Raz fait.⁵⁰

Une autre version modérée de l'approche évaluative est celle développée par Katja Maria Vogt dans son article intitulé « Desiring the Good : A Socratic Reading of Aristotle ». La thèse de Vogt s'inspire d'une idée aristotélicienne qui affirme que la raison ultime de chacune de nos actions est de s'assurer une vie bonne. Par exemple, je vais au marché parce que je veux acheter de la nourriture. Si je me demande pourquoi je veux acheter de la nourriture, la raison en est que je veux manger. Je veux manger parce que je veux que ma vie aille bien. Donc, ultimement, je vais au marché parce que je veux que ma vie aille

⁴⁷ *Ibid.*, pp. 43-44; Raz, 2008, pp. 17-19.

⁴⁸ Raz, 1999, p. 32.

⁴⁹ *Ibid.*, pp. 33-34.

⁵⁰ Raz, 2008, pp. 16-18.

bien.⁵¹ Ainsi, selon Vogt, nous avons une motivation d'arrière-plan (*background motivation*) en faveur de la vie bonne et cette motivation générale donne de la force à nos motivations particulières. Si je suis motivé à aller au marché, c'est parce que cela va contribuer au fait que ma vie aille bien et que j'ai une motivation supérieure qui me pousse à rechercher la vie bonne.

L'approche évaluative s'applique mal aux actions particulières, selon Vogt. Elle doit plutôt être une théorie à propos de la motivation en général.⁵² Si notre motivation d'arrière-plan en faveur de la vie bonne guide certaines de nos actions particulières, les actions triviales échappent souvent à cette influence. Par exemple, si je me demande quel biscuit choisir dans une boîte, il est peu probable que cela demande un véritable effort de réflexion quant à la valeur de ce choix dans ma vie. De nombreuses actions de la vie de tous les jours ne nous semblent pas assez importantes pour mériter un tel effort de délibération.⁵³ Pour cette raison, l'approche évaluative n'explique pas efficacement les actions particulières et laisse la porte ouverte à de nombreuses exceptions. Les actions insignifiantes et certaines mauvaises actions de moindre importance passent en quelque sorte à travers le filtre de la motivation d'arrière-plan et ne peuvent pas être expliquées par l'approche évaluative.

La motivation d'arrière-plan en faveur de la vie bonne n'est pas non plus infaillible. Elle peut être affectée, affaiblie ou même détruite par les aléas de la vie. Contrairement à ce que croyait Aristote, il est indéniable, pense Vogt, que certaines personnes ne veulent pas que leur vie aille bien.⁵⁴ Lorsque la motivation d'arrière-plan est affectée par la dépression ou par d'autres affections, une personne perd le goût de rechercher la vie bonne, elle perd ses repères par rapport aux valeurs qu'elle croyait acquises. La perte de cette motivation peut expliquer certains cas extrêmes d'acrasie ou encore le fait qu'une personne puisse désirer le mal. Dans la plupart des cas, les gens ont naturellement tendance à rechercher le bien dans leur vie. Toutefois, cela ne forme pas un principe

⁵¹ Vogt, 2010, pp. 7-8.

⁵² *Ibid.*, p. 22.

⁵³ *Ibid.*, p. 23.

⁵⁴ *Ibid.*, pp. 24-25.

infaillible, selon Vogt, et la motivation en faveur de la vie bonne peut être ruinée par des événements dans la vie d'une personne.

La thèse de Vogt est moins forte, c'est-à-dire moins contraignante, que celle de Raz. La thèse de Raz concerne les actions particulières, tandis que celle de Vogt concerne une certaine motivation d'arrière-plan présente dans notre vie. La thèse de Raz postule un lien conceptuel, infaillible, entre raison et valeur. La thèse de Vogt, pour sa part, fait un constat empirique en affirmant que généralement, les gens désirent que leur vie aille bien. Pour Raz, les exceptions à l'approche évaluative sont soit des cas dégénérés, anomiques, ou encore des types d'action fondamentalement différents de ceux expliqués par la théorie. Vogt, pour sa part, affirme que les actes anomiques sont simplement le résultat normal de circonstances différentes. Si notre motivation d'arrière-plan est affaiblie ou détruite, alors il est naturel de désirer le mal ou d'être victime d'acrasie. Bien que les thèses de Raz et de Vogt soient toutes les deux des versions modérées de l'approche évaluative lorsqu'on les compare à la thèse de Tenenbaum, l'approche de Raz est bien plus radicale que celle de Vogt. Si Raz et Tenenbaum postulent tous les deux certains principes fondamentaux concernant l'action et l'intention humaines, Vogt, pour sa part, se contente de postuler une observation concernant la psychologie humaine en général. Pour cette raison, l'approche de Vogt est moins pertinente pour notre recherche. Nous nous concentrerons donc davantage sur les thèses de Tenenbaum et Raz dans le reste de ce travail.

CHAPITRE II

OBJECTIONS À L'APPROCHE ÉVALUATIVE

2.1 L'argument des contre-exemples

L'objection la plus courante adressée à l'approche évaluative consiste à formuler des contre-exemples allant à l'encontre de la théorie. La plupart des opposants à l'approche évaluative pensent qu'il est possible de présenter un certain nombre de cas qui contredisent les principes mis de l'avant par la théorie. Ces contre-exemples sont censés réfuter l'approche évaluative en démontrant que dans certains cas, la perception du bien peut échouer à motiver et que, dans d'autres cas, autre chose que le bien, par exemple le mal, peut motiver un individu à agir. Michael Stocker offre un exemple paradigmatique de ce genre d'objections dans son article intitulé « Desiring the Bad : An Essay in Moral Philosophy ». ⁵⁵ Dans cet article, Stocker formule l'hypothèse selon laquelle désirer le mal et ne pas désirer le bien sont des phénomènes courants dans notre vie et que cela met en évidence le fait qu'il n'existe pas de relation simple entre motivation et évaluation, et que par conséquent l'approche évaluative est fautive.

La stratégie argumentative de Stocker s'appuie sur la présentation de contre-exemples pouvant réfuter l'approche évaluative. Pour être en mesure de prouver qu'il est possible d'être motivé par le mal et de ne pas être motivé par le bien, il suffit de démontrer qu'il existe des cas évidents où le fait que l'action soit perçue comme étant mauvaise constitue un élément motivateur pour l'agent et que, parallèlement, dans certains cas la perception par l'agent qu'une action est bonne ne suffit pas à le motiver à accomplir cette action. ⁵⁶ Si Stocker arrive à nous convaincre que des contre-exemples crédibles existent, il aura réussi à démontrer que le bien ne motive pas nécessairement et que d'autres choses, comme le mal, peuvent motiver. Il aura donc réussi à réfuter l'approche évaluative, ou à tout le moins la version radicale de la théorie.

⁵⁵ Stocker, 1979, pp. 738-739.

⁵⁶ *Ibid.*, p. 741.

Le premier contre-exemple invoqué par Stocker a pour but de démontrer que le bien peut échouer à motiver. C'est le cas d'un politicien qui faisait de son mieux, dans sa jeunesse, pour aider les gens de sa communauté.⁵⁷ En vieillissant, il a perdu le désir d'aider des gens autres que sa famille et ses proches. Il sait qu'il pourrait encore faire beaucoup de bien aux gens de sa communauté s'il s'en donnait la peine, mais il n'en a plus envie. Il est aigri et ne se soucie plus des gens qu'il souhaitait autrefois aider. Dans cet exemple, le politicien est conscient du bien qu'il pourrait accomplir en faisant l'effort d'aider sa communauté, mais il n'est pas motivé par ce bien. Il préfère ne pas se soucier des autres et les laisser se débrouiller seuls. Pour Stocker, l'attitude du politicien réfute l'approche évaluative parce qu'elle démontre que le bien échoue à le motiver. Cet exemple est un cas typique d'acédie où un agent sait qu'une action est bonne, mais n'est pas motivé à l'accomplir.

On pourrait répondre à Stocker que le politicien est encore motivé par le bien même lorsqu'il refuse d'aider sa communauté. Il est possible que ce qui le motive soit la paix d'esprit qu'il souhaite avoir en ne se mêlant plus de projets politiques angoissants, ou qu'il préfère favoriser davantage son propre bien et celui de ses proches plutôt que celui d'étrangers. Si tel est le cas, le politicien est encore motivé par le bien, sauf qu'il s'agit d'un bien égoïste qui consiste à faire uniquement ce qui est bon pour lui-même et pour ses proches sans se soucier de la communauté. Si on accepte cette explication, alors on peut reprocher au politicien son égoïsme, mais on doit tout de même reconnaître qu'il agit en vue du bien.

Stocker est en désaccord avec cette interprétation. Pour lui, le fait d'être aigri et de ne pas se soucier de sa communauté est une explication suffisante pour comprendre le comportement du politicien.⁵⁸ Il n'est pas besoin d'invoquer des biens rivaux et incompatibles pour expliquer le fait que le politicien n'aide pas sa communauté. Il est aigri et n'a aucun désir de faire ce qui est bien pour sa communauté, tout simplement. Toute autre explication est superflue. Stocker pense que plusieurs facteurs tels que

⁵⁷ *Ibid.*, pp. 741-742.

⁵⁸ *Ibid.*, p. 743-744.

l'humeur, l'énergie et l'intérêt peuvent expliquer pourquoi, dans certains cas, nous ne sommes pas motivés à accomplir ce que nous croyons être bien. Lorsque nous sommes dans un état de fatigue, de maladie, d'apathie, de désespoir ou de honte, il est courant de ne pas être motivé à rechercher le bien, tant le nôtre que celui des autres. Dans ces situations, nous sommes en mesure de percevoir le bien qui pourrait découler de nos actions, sauf que celui-ci échoue à nous motiver. Ainsi, conclut Stocker, il a été démontré avec cet exemple que le bien peut échouer à motiver.⁵⁹ Le premier pilier de l'approche évaluative, qui dit que le bien motive à agir, a donc été réfuté.

Stocker souhaite également démontrer que d'autres considérations que le bien peuvent motiver l'individu à agir. Il formule l'hypothèse qu'il est possible de désirer le mal et que le mal en soi peut constituer un facteur motivant l'individu à agir.⁶⁰ Pour étayer cette hypothèse, il introduit l'idée selon laquelle nous souhaitons parfois faire du mal à autrui. Confrontés à cette situation, les partisans de l'approche évaluative font face à un dilemme. Pour éviter de dire que le mal motive, ils doivent affirmer, selon Stocker, soit que le mal causé à autrui est perçu comme un bien par l'agent, soit que l'action de faire du mal à autrui est toujours effectuée en vue d'un but différent et que le mal en soi n'est pas l'objet du désir.⁶¹ Stocker affirme qu'il n'est pas plausible de penser que l'acte de faire du mal à autrui puisse être perçu comme bon en soi. Il ne voit aucune raison non plus de penser que chaque action nocive doive nécessairement être perpétrée en vue d'un bien quelconque qui soit le véritable objet du désir. Lorsque nous sommes en colère, jaloux, rejetés, frustrés ou menacés, causer du mal à autrui peut être précisément ce que nous désirons. Il est aussi possible de vouloir se faire du mal à soi-même lorsque nous sommes dans certains états d'esprit autodestructeurs tels que la honte ou la culpabilité. À la lumière de ces arguments, Stocker conclut que nous ne désirons pas seulement ce qui est bien; nous pouvons aussi désirer ce qui est mal.

⁵⁹ *Ibid.*, p. 745.

⁶⁰ *Ibid.*, pp. 746-747.

⁶¹ *Ibid.*, p. 748.

David Velleman renchérit sur ce point dans son article intitulé « The Guise of the Good ». ⁶² Il avance l'argument que les désirs et les raisons d'agir d'un agent peuvent être perversis, c'est-à-dire orientés vers le mal plutôt que vers le bien. Il est possible, selon lui, de désirer quelque chose justement parce qu'il s'agit de quelque chose de mauvais. Afin d'illustrer son point de vue, Velleman utilise l'exemple d'une personne dans un état de dépression et de désespoir. Il n'est pas rare qu'une personne dans un tel état ressente du dégoût à l'idée de faire ce qui est bon pour elle-même et pour les autres. Dans l'exemple de Velleman, cette personne prend la résolution de ne plus jamais faire d'efforts en vue du bien et de s'abstenir d'accomplir toute action qui pourrait lui sembler bonne. Si elle perçoit de la valeur dans une action, elle évite à tout prix de l'accomplir. ⁶³ Dans cet exemple, c'est le désespoir, plutôt qu'un désir de rechercher le bien, qui motive cette personne. Celle-ci agit sous le coup du désespoir et non pas en raison du bien qu'elle peut percevoir. Ses motivations perverses font en sorte que ses actions ne sont plus compatibles avec l'approche évaluative. Velleman s'accorde donc avec Stocker pour dire que l'argument des contre-exemples réussit à réfuter l'approche évaluative puisqu'il pense lui aussi qu'il est possible d'être motivé non pas par le bien, mais plutôt par le mal.

Contrairement à d'autres, Stocker ne voit pas les contre-exemples invoqués contre l'approche évaluative comme des cas anormaux ou des exceptions. Pour lui, ces cas démontrent tout simplement que la relation directe entre évaluation et motivation postulée par les partisans de l'approche évaluative n'existe pas. ⁶⁴ Les contre-exemples ne sont pas des exceptions à la règle puisqu'une telle règle n'existe tout simplement pas. Stocker considère que l'esprit humain est beaucoup plus complexe que ce que les théories philosophiques traditionnelles semblent le suggérer. Plusieurs facteurs tels que l'énergie, l'humeur et l'intérêt ont une influence aussi bien sur la motivation d'une personne que sur son évaluation de ce qui est bon ou mauvais. L'évaluation et la motivation ne sont reliées entre elles que par l'intermédiaire de structures psychologiques complexes et la philosophie moderne doit pouvoir reconnaître cette complexité.

⁶² Velleman, 1992, p. 20.

⁶³ *Ibid.*, p. 21.

⁶⁴ Stocker, 1979, pp. 738-739, 749-752.

Même si l'argument des contre-exemples peut sembler convaincant, sa capacité à réfuter entièrement l'approche évaluative est controversée même chez les détracteurs de la théorie. Kieran Setiya, bien qu'il s'oppose à l'approche évaluative et qu'il reconnaisse la validité des contre-exemples, admet que l'argument des contre-exemples ne peut pas réussir à lui seul à réfuter l'approche évaluative. Selon lui, les partisans de la théorie trouveront toujours un moyen d'interpréter les contre-exemples différemment afin de démontrer qu'une action apparemment mauvaise a été faite en vue du bien.⁶⁵

Il existe deux stratégies argumentatives possibles pour contrer l'argument des contre-exemples. La première consiste à réduire la portée de la théorie afin d'accommoder les contre-exemples et d'en faire des exceptions à l'approche évaluative. Si on adopte cette stratégie, on peut reconnaître que certains contre-exemples sont valides sans toutefois être forcé d'admettre qu'ils réfutent la théorie. La seconde stratégie consiste à rejeter la validité des contre-exemples et à les interpréter différemment pour qu'ils s'accordent avec l'approche évaluative. Raz adopte principalement la première stratégie, et Tenenbaum la seconde.

La théorie de Raz est bâtie sur mesure pour répondre à l'argument des contre-exemples. Comme nous l'avons vu précédemment, Raz admet que certaines actions triviales puissent être faites sans raison, se retrouvant ainsi hors de la gouverne de l'approche évaluative. Il admet aussi que les raisons d'agir peuvent être perverties ou anomiques et que, lorsqu'il en est ainsi, une action peut être accomplie en vue du mal plutôt que du bien. Les actions faites sous le coup de l'émotion peuvent aussi constituer des exceptions à la théorie. Cependant, bien que Raz admette l'existence de ces cas exceptionnels, il se montre intransigent à d'autres égards. Par exemple, il refuse de concéder qu'une raison d'agir puisse être véritablement neutre du point de vue de la valeur.⁶⁶ Une raison d'agir présente typiquement l'action comme étant bonne ou désirable. Dans certains cas exceptionnels, le fait qu'une action soit mauvaise peut constituer une raison psychologique d'agir. Dans les deux cas, le concept de valeur entre en jeu. Raz refuserait

⁶⁵ Setiya, 2007, p. 38; Setiya, 2010, p. 83.

⁶⁶ Raz, 1999, pp. 30-34.

donc un contre-exemple qui servirait à affirmer que l'on peut avoir une raison de faire une chose sans que cette chose ne nous apparaisse comme bonne (ou mauvaise).

Il est difficile également de réconcilier la thèse de Raz avec les cas d'acrasie ou d'acédie présentés par les opposants à l'approche évaluative. Avec l'exemple du politicien, par exemple, Raz devrait affirmer que si le politicien pense qu'il a une raison de ne pas aider sa communauté, il doit percevoir cette action (ou cette inaction) comme ayant une certaine valeur, et qu'il recherche un bien incompatible tel que la paix d'esprit. On pourrait aussi interpréter l'exemple différemment et dire que c'est sous le coup d'une émotion, par exemple la frustration, qu'il agit ainsi, ou encore que le politicien est fondamentalement méchant et qu'il agit pour une raison anémique, qui est le mal et la souffrance de la communauté. Raz serait moins conciliant, par contre, envers l'idée que le politicien perçoit véritablement le fait d'aider sa communauté comme un bien, mais qu'il n'est pas motivé à agir ainsi. Raz admet donc, jusqu'à un certain point, des exceptions à l'approche évaluative, comme les actions faites sans raison et les actions faites en vue du mal, mais il se montre intransigeant face à l'idée que des raisons neutres du point de vue de la valeur puissent motiver, et envers les cas typiques d'acrasie et d'acédie.

Tenenbaum, pour sa part, fait face à un défi plus considérable s'il souhaite récuser l'argument des contre-exemples. Il doit démontrer, dans la mesure du possible, que les contre-exemples proposés par les détracteurs de la théorie ne sont pas d'authentiques exceptions et que même dans ces cas-là, l'agent agit *sub specie boni*, c'est-à-dire en vue du bien. Tenenbaum s'attaque à trois types de contre-exemples : les actions faites en vue du mal, les actions acratiques et les cas d'acédie.

Tenenbaum affirme qu'il est impossible d'agir strictement en vue du mal. Autrement dit, les aspects purement négatifs d'une action ne peuvent pas contribuer à rendre cette action désirable à nos yeux et nous motiver à l'accomplir. Tout ce qui est désiré doit nécessairement être perçu comme bon d'une certaine manière. Il n'y a pas d'inversion de valeur possible, le mal ne motive pas.⁶⁷ Mais comment expliquer alors que certaines

⁶⁷ Tenenbaum, 2007, p. 14.

personnes commettent des actions mauvaises? Certains auteurs utilisent le diable comme exemple hypothétique pour illustrer ce point.⁶⁸ S'il existe un personnage qui agit en vue du mal, disent-ils, ce doit être le diable. Comment peut-on affirmer, alors, que le diable est motivé par le bien et non par le mal? Pour répondre à cette question, il faut concevoir un être qui désire ce qui est moralement mauvais, qui fait du mal moral son bien *simpliciter*.⁶⁹ Ainsi, lorsque le diable fait du mal à autrui, il accomplit ce qui est bon pour lui-même. C'est peut-être la seule façon de comprendre la motivation d'un être diabolique, puisque si cet agent désirait vraiment le mal *simpliciter*, il serait incapable d'agir de manière cohérente. Si le diable était uniquement motivé par ce qui lui semble mauvais, alors il serait incapable de planifier ses actions, car l'acte de planification est quelque chose d'utile et de désirable, donc de bon.⁷⁰ Il rejetterait toute cohérence et refuserait de faire quoi que ce soit qui puisse l'aider et lui être utile. Un tel diable serait alors complètement impuissant à faire quoi que ce soit. Pourtant, personne ne pense que le diable se comporterait de cette manière s'il existait. Au contraire, lorsqu'on s'imagine le personnage du diable, on pense à un être puissant qui désire faire du mal aux autres, qui prend plaisir à les voir souffrir et qui trouve son propre bonheur dans le malheur d'autrui. Il faut donc concevoir le diable comme un être motivé par le bien *simpliciter*, mais qui perçoit le mal moral comme faisant partie de ce qui est bon pour lui. Ainsi, selon Tenenbaum, le mal ne peut pas réellement motiver, et par conséquent les contre-exemples présentant des actions faites en vue du mal ne peuvent pas être valides, puisque même lorsque nous accomplissons des actions immorales, nous percevons ces actions comme désirables d'un certain point de vue.

Tenenbaum s'attaque ensuite aux cas d'acrasie. Lorsqu'une personne est victime d'acrasie, cela signifie qu'elle sait qu'une certaine action est bonne et qu'elle devrait accomplir cette action, mais qu'elle manque tout simplement de motivation à agir. On peut imaginer une personne qui sait qu'elle devrait sortir pour faire les courses, mais qui choisit de rester chez elle naviguer sur Internet. Les adversaires de l'approche évaluative utiliseraient un exemple comme celui-ci pour dire que le bien peut échouer à motiver et

⁶⁸ Anscombe, 1963, p. 75; Tenenbaum, 2007, pp. 253-256; Velleman, 1992, pp. 18-19.

⁶⁹ Tenenbaum, 2007, p. 253.

⁷⁰ *Ibid.*, p. 255.

donc que l'approche évaluative est fautive. Tenenbaum affirme au contraire que même dans les cas qui semblent acratiques, l'agent finit toujours par faire ce qui lui semble le mieux, et que par conséquent l'acrasie véritable n'existe pas. Il s'appuie sur un cas célèbre dans l'histoire de la philosophie : les *Méditations* de René Descartes.⁷¹ Lorsque Descartes est plongé dans ses méditations, il est en mesure de comprendre qu'il vaut mieux douter de tout ce qu'il perçoit, car le monde n'est peut-être qu'une illusion créée par un malin génie. Toutefois, lorsqu'il n'est pas en train de méditer et qu'il vit sa vie de tous les jours, il ne peut pas s'empêcher de croire que le monde qui l'entoure est réel, et il agit en fonction de cette impression même s'il sait, en son for intérieur, qu'il devrait douter de tout.

Tenenbaum introduit la distinction entre compréhension directe et compréhension oblique pour expliquer ce phénomène. Nous avons une compréhension directe d'une idée ou d'un concept lorsqu'on est directement confronté à la vérité de cette idée ou de ce concept. Quand Descartes est en pleine méditation, son raisonnement le mène à conclure qu'il doit douter de tout. Dans ce moment de contemplation méditative, il a une compréhension directe du monde en tant qu'illusion. Par contre, lorsqu'il n'est pas en train de méditer et qu'il interagit avec le monde, l'idée que le monde est une illusion ne lui apparaît plus comme une vérité évidente, mais plutôt comme une simple notion abstraite. Il a maintenant une compréhension oblique de cette idée. Il se la remémore, il sait qu'il a précédemment jugé que cette idée est vraie, mais elle ne lui apparaît plus comme une évidence. Dans cette situation, ce qui lui apparaît évident, c'est que le monde physique est réel, et il agit en conséquence. Les deux idées contradictoires, que le monde est réel et qu'il est une illusion, entrent en conflit dans son esprit, mais dans la vie de tous les jours, c'est la compréhension directe du monde en tant que réalité qui l'emporte sur la compréhension oblique du monde en tant qu'illusion. Il est possible qu'une compréhension oblique l'emporte sur une compréhension directe dans certains cas, mais dans ce cas-ci c'est la compréhension directe qui l'emporte.

⁷¹ *Ibid.*, pp. 268-274.

Dans les cas d'acrasie, il se produit un conflit similaire entre une compréhension directe qu'une action est bonne et une compréhension oblique favorisant une action différente.⁷² Prenons l'exemple d'une personne qui sait qu'elle devrait accomplir une certaine corvée, mais qui choisit de s'amuser à la place. C'est un cas classique de procrastination. L'agent comprend de manière oblique qu'accomplir la corvée est une bonne chose, mais la corvée ne lui apparaît pas comme une option attrayante. L'idée de ne pas exécuter la corvée et de s'amuser à la place est une option beaucoup plus plaisante. L'agent a une compréhension directe du bien de s'amuser et il désire s'amuser plutôt que travailler. Nous avons ici deux compréhensions contradictoires du bien qui entrent en conflit dans l'esprit de l'agent. Si la compréhension oblique de la valeur de l'accomplissement de la corvée échoue à s'imposer face à la compréhension directe du bien de s'amuser, on dit alors que l'agent souffre de faiblesse de la volonté. Par contre, il serait faux de dire que c'est un véritable cas d'acrasie où l'agent manque de motivation à accomplir ce qui lui semble bon. Il s'agit plutôt d'un cas où l'agent fait ce qui lui semble être la meilleure chose à faire même si ce n'est pas ce que la raison lui dicterait.⁷³ Si l'agent n'est pas suffisamment motivé à accomplir sa corvée, c'est qu'il n'est pas pleinement convaincu qu'il s'agit de la bonne chose à faire. Ultimement, il a plus envie de s'amuser. Les cas que l'on considère généralement comme étant des cas d'acrasie sont donc, en fin de compte, des cas où l'agent est motivé à faire ce qui lui semble bien. Il n'y a pas de véritables cas d'acrasie selon Tenenbaum.

Tenenbaum rejette finalement les cas d'acédie (*accidie*) censés démontrer que le bien peut totalement échouer à motiver et qu'un agent peut croire que X est bon sans ressentir la moindre motivation à faire X. Pour y parvenir, Tenenbaum introduit le concept de conditionnalité (*conditioning*).⁷⁴ Il est possible pour un agent de penser qu'une action X est bonne en soi, mais que sous certaines conditions, elle ne serait pas désirable. Pour illustrer cette idée, Tenenbaum propose un exemple dans lequel Isabel pense que faire un voyage à New York est une bonne chose en soi, mais qu'elle ne serait pas motivée à faire

⁷² *Ibid.*, p. 274.

⁷³ *Ibid.*, p. 282.

⁷⁴ *Ibid.*, p. 286.

ce voyage sous certaines conditions.⁷⁵ En effet, si la seule manière de pouvoir faire ce voyage était de trahir la confiance d'un ami, alors cet acte immoral éliminerait tout plaisir subséquent qui pourrait découler du voyage à New York. Isabel peut penser que le voyage en soi est une bonne chose, mais si ce voyage est rendu possible par un acte immoral, alors il devient indésirable parce que le sentiment de culpabilité ressenti par Isabel pendant le voyage l'empêcherait d'en profiter. Nous avons donc un exemple qui pourrait passer pour un cas d'acédie, mais en réalité, lorsqu'on examine la situation dans son ensemble, on reconnaît qu'un voyage rendu possible par un acte de trahison n'est pas quelque chose de désirable. Certaines conditions doivent être remplies pour qu'Isabel soit motivée à faire ce voyage. Une de ces conditions est que le voyage lui-même ne soit pas rendu possible par un acte immoral. Si cette condition n'est pas remplie, alors elle n'est pas motivée à faire le voyage puisque la situation dans son ensemble ne lui apparaît pas comme bonne.

Tenenbaum pense qu'en utilisant la notion de conditionnalité, on peut également expliquer les cas de dépression sévère où une personne cesse d'être motivée par son propre bien.⁷⁶ On peut postuler que certaines conditions, comme l'estime de soi, doivent être remplies pour qu'une personne désire son propre bien, mais que parfois elles ne le sont pas. Tant que ces conditions ne sont pas remplies, l'agent ne perçoit pas l'idée de rechercher son propre bien comme quelque chose de désirable, toutes choses considérées. Cette notion de conditionnalité peut nous permettre de réinterpréter de manière plausible les cas supposés d'acédie afin de démontrer que même dans ces cas-là, l'agent agit en conformité avec l'approche évaluative.

S'il se montre intransigeant face aux cas supposés d'acrasie, d'acédie et d'actions faites en vue du mal, Tenenbaum est prêt à admettre certaines actions pathologiques comme de véritables exceptions à l'approche évaluative. Si un agent accomplit compulsivement une action, comme quelqu'un qui se sentirait obligé de constamment se laver les mains, alors

⁷⁵ *Ibid.*, pp. 287-289.

⁷⁶ *Ibid.*, pp. 293-294.

l'action n'est pas accomplie *sub specie boni* même s'il s'agit d'une action volontaire.⁷⁷ En effet, si on demande à l'agent pourquoi il accomplit cette action, il répond qu'il ne sait pas pourquoi, mais qu'il ne peut pas s'en empêcher. L'action compulsive ne répond donc pas au critère du « pourquoi? » d'Anscombe, et n'est pas considérée comme une action intelligible. Elle n'est donc pas explicable par l'approche évaluative.

Que penser, en fin de compte, de l'argument des contre-exemples proposé par Stocker et Velleman? Pour ma part, je pense que la conception du bien de Stocker est une conception partiellement morale et que cela affecte la validité de son argument. Dans son article « Desiring the Bad », Stocker affirme qu'il est invraisemblable qu'une personne puisse penser que l'acte de blesser quelqu'un est une bonne chose. Cette affirmation l'aide ensuite à réfuter l'approche évaluative.⁷⁸ Or, il m'apparaît clairement que ceci est une affirmation morale et qu'il n'y a aucune raison de penser que le fait de blesser autrui ne puisse jamais apparaître à quiconque comme une chose bonne ou désirable. Le contraire me semble plus plausible. Une théorie de l'action crédible doit être en mesure d'expliquer pourquoi certaines personnes commettent des actes abominables. Bien qu'il puisse être réconfortant de penser que personne ne pourrait concevoir l'acte de blesser ou tuer comme quelque chose de bon, cette explication ne nous aide pas à comprendre le phénomène. Si les auteurs de crimes atroces percevaient véritablement leurs propres actions comme répugnantes et ne leur accordaient aucune valeur, nous serions alors bien en peine d'expliquer pourquoi ils les commettent. Ils seraient, comme l'immense majorité d'entre nous, révoltés à l'idée de faire du mal à un autre être humain et se comporteraient donc de manière décente. Pourtant, ils ne sont pas comme l'immense majorité d'entre nous puisqu'ils commettent bel et bien de tels actes. Nous avons vu, avec l'exemple du diable, qu'il est problématique, voire impossible d'imaginer quelqu'un agissant en vue du mal *simpliciter*. Nous devons donc conclure qu'il est non seulement possible de percevoir l'acte de blesser autrui comme quelque chose de désirable, mais que les gens qui commettent des crimes atroces perçoivent leurs propres actions de cette manière au moment où ils les commettent.

⁷⁷ *Ibid.*, pp. 236-238.

⁷⁸ Stocker, 1979, p. 748.

Afin de démontrer que le bien peut échouer à motiver, Velleman utilise l'exemple de quelqu'un qui prendrait consciemment la résolution de ne plus faire ce qui lui semble bon.⁷⁹ Si quelque chose lui semble bon, il doit s'en abstenir. Une telle attitude n'est pas seulement impossible dans les faits, elle génère également un paradoxe. Si une personne refuse de faire ce qui lui semble bon, c'est qu'elle préfère être cohérente avec sa résolution de ne plus faire ce qui lui semble bon. Le problème, c'est que le désir d'être cohérent renvoie lui aussi à un jugement de valeur qu'il est bon de respecter les résolutions que l'on prend. L'agent ne peut pas s'extirper de cette situation parce que chacune de ses décisions renvoie ultimement à un jugement sur ce qu'il vaut mieux faire, donc à la recherche d'une valeur. Dans tous les cas, il continue de faire ce qui lui semble le mieux dans sa situation actuelle.

Un autre argument avancé par Velleman est qu'aux yeux d'une personne désespérée, l'idée de faire ce qui est bon pour soi et d'améliorer sa condition peut apparaître comme une illusion absurde, et que cela constitue un exemple où le bien échoue à motiver.⁸⁰ Pourtant, si une chose apparaît comme une illusion absurde, elle n'apparaît pas comme valable. Velleman fait l'erreur de penser que ce qu'une personne croit être bon, lorsqu'elle est dans son état normal, peut se transposer sur un état d'esprit différent, comme le désespoir. Il suppose que si une personne désespérée ne désire pas faire les choses qui lui sembleraient bonnes dans son état normal, c'est qu'elle n'est pas motivée par le bien. Pourtant, il est plus naturel de penser que lorsque notre état d'esprit change, notre perception du bien change également. Pour la personne désespérée, il peut sembler bon de rester couchée et l'idée de faire un effort pour améliorer sa condition peut sembler répugnante. Si tel est le cas, alors les actions de cette personne ne constituent pas un véritable contre-exemple à l'approche évaluative. Par ailleurs, les autres contre-exemples de Stocker et Velleman se fondent également sur des états d'esprit anormaux. Le politicien est aigri, l'agresseur est furieux, la personne acédique est désespérée.⁸¹ Je ne crois pas que lorsque nous sommes dans de tels états d'esprit, le principe selon lequel

⁷⁹ Velleman, 1992, p. 21.

⁸⁰ *Ibid.*, p. 20.

⁸¹ Stocker, 1979, pp. 738-739, 748; Velleman, 1992, pp. 20-21.

nous sommes motivés à faire ce qui nous semble bon disparaît ou est affecté. C'est plutôt notre perception du bien qui change. Dans chaque instance, nous sommes motivés par ce qui nous apparaît bon, mais quand notre état d'esprit est altéré par l'émotion ou par d'autres facteurs, nous percevons différentes choses comme étant bonnes et nous sommes motivés à faire ce qui, dans notre état d'esprit normal, pourrait nous sembler mauvais ou inutile.

Stocker affirme que l'esprit humain est bien plus complexe que ne le laisse entendre l'approche évaluative et qu'il n'existe pas de lien direct entre la perception du bien et la motivation à agir, mais plutôt que la motivation est gouvernée par une structure complexe au sein de laquelle divers éléments tels que l'humeur, l'intérêt et l'énergie ont une influence.⁸² Il est indéniable que quelque chose comme l'humeur peut avoir une influence sur notre motivation et sur la manière dont nous agissons, mais l'humeur affecte-t-elle le principe que nous sommes motivés par ce qui nous semble bon? Si l'on considère que ce qui nous apparaît bon à un moment donné est une variable, et que cette variable est déterminée entre autres par les éléments mentionnés par Stocker, par exemple l'humeur, l'intérêt et l'énergie, alors on peut tenir compte de la complexité de l'esprit humain tout en préservant le principe que le bien motive. Si différents facteurs ont une influence sur notre comportement, c'est parce que ces facteurs ont une influence sur notre perception du bien, et que par extension notre perception du bien détermine notre motivation à agir. L'erreur que commettent Stocker et Velleman est de penser que notre conception du bien est immuable et purement rationnelle alors qu'elle est en réalité fluide et change selon les conditions dans lesquelles on se trouve.

Même si les apparences de bien peuvent être fluides et changer selon les circonstances, on peut difficilement rejeter l'idée que nos actions sont toujours motivées par une considération pour le bien. Il semble que peu importe l'action intentionnelle, si on l'analyse en profondeur, on revient toujours en fin de compte à considérer en quoi elle est bonne. Revisitons l'exemple utilisé par Katja Maria Vogt dans « *Desiring the Good : A Socratic Reading of Aristotle* », celui d'une personne qui va au marché pour acheter de la

⁸² Stocker, 1979; pp. 738-739.

nourriture.⁸³ Pourquoi achète-t-elle de la nourriture? Pour pouvoir manger. Pourquoi veut-elle manger? Pour ne pas souffrir de la faim. Pourquoi vouloir éviter de souffrir de la faim? Parce que c'est bon. Une fois que l'on en vient à invoquer le bien comme raison d'agir, on ne peut plus pousser le raisonnement plus loin. Si on lui demande pourquoi désirer ce qui est bon, la réponse est encore une fois : parce que c'est bon. Ainsi, puisque l'analyse de nos propres raisonnements ramène aussi infailliblement vers la considération du bien, il devient alors difficile de rejeter l'approche évaluative et nier qu'il existe un lien direct entre évaluation et motivation.

En conclusion, l'argument des contre-exemples tel que présenté par Stocker et Velleman échoue à convaincre de la fausseté de l'approche évaluative. Les détracteurs de la théorie accusent ses défenseurs de vouloir réinterpréter à tout prix les contre-exemples afin de sauver l'approche évaluative, mais en réalité, les réinterprétations offertes par les partisans de la théorie offrent souvent une explication plus plausible de ces situations que l'idée qu'elles ne sont pas des cas où l'agent agit *sub specie boni*. La distinction entre compréhension directe et oblique apportée par Tenenbaum est particulièrement judicieuse pour expliquer les cas supposés d'acrasie. Jusqu'à présent, nous n'avons pas encore rencontré un seul contre-exemple qui soit à la fois réaliste et sans équivoque. Il est douteux que de tels cas existent réellement et il revient aux détracteurs de l'approche évaluative de prouver le contraire. D'ici là, nous devons conclure que l'argument des contre-exemples échoue à mettre en danger la théorie.

2.2 L'argument de la sophistication intellectuelle

Le deuxième argument contre l'approche évaluative que nous étudierons dans ce travail est présenté par David Velleman dans son article « The Guise of the Good ». Il s'agit de l'argument de la sophistication intellectuelle. Afin de bien comprendre la pertinence de cet argument, nous devons le présenter dans son contexte. David Velleman considère que l'approche évaluative est une tentative de former un hybride de deux théories de l'action :

⁸³ Vogt, 2010, pp. 7-8.

la théorie de la motivation et la théorie de la direction rationnelle (*rational guidance*).⁸⁴ Ceux qui adoptent la théorie de la motivation, selon Velleman, croient que les individus agissent parce qu'ils ont un désir et la croyance qu'une action particulière comblera ce désir. L'action est expliquée par le fait que l'agent a une certaine prédisposition favorable envers un certain état de fait, tel que connaître l'heure, et que cette prédisposition le pousse à regarder sa montre afin de faire survenir cet état de fait.⁸⁵ Cette explication laisse de côté des notions telles que la rationalité et la justification de l'action. Elle fait appel à une sorte de force brute, qui est le désir, sans expliquer en quoi celui-ci peut être justifié ou inclure un jugement de valeur envers la chose désirée. Rien n'explique en quoi le fait de connaître l'heure peut constituer quelque chose de bon ou de valable pour l'agent; il est seulement tenu pour acquis que l'agent possède une certaine disposition favorable envers la possibilité de connaître l'heure.

La théorie de la direction rationnelle, pour sa part, affirme que dans une certaine proposition comprise par l'agent se trouve une force justificatrice intrinsèque qui amène l'agent à accomplir l'action à condition qu'il croie en la vérité de cette proposition.⁸⁶ L'agent pourrait croire, par exemple, la proposition suivante : « Il est justifié pour moi de regarder ma montre afin de connaître l'heure ». En acceptant la vérité de cette proposition, il accepte la rationalité de l'action et il est mû par la force justificatrice de la proposition. Ainsi, l'action de regarder sa montre devient en quelque sorte la conclusion d'une inférence logique dans laquelle la vérité des prémisses garantit la vérité de la conclusion. Dans cette situation, l'agent n'exprime pas un certain biais en faveur d'une proposition plutôt qu'une autre; il se contente de tenir pour vraie une proposition qui contient en elle-même une justification en faveur de l'action.

L'approche évaluative, selon Velleman, est une tentative de réconcilier ces deux théories de l'action et d'ajouter un élément cognitif à la motivation afin de lui donner un aspect plus rationnel.⁸⁷ Selon cette nouvelle théorie, un contenu justifiant l'action se trouve dans

⁸⁴ Velleman, 1992, p. 3.

⁸⁵ *Ibid.*, pp. 3-4.

⁸⁶ *Ibid.*, p. 4.

⁸⁷ *Ibid.*, p. 6.

l'attitude motivationnelle elle-même. Autrement dit, les désirs sont des jugements de valeur possédant une force justificatrice.⁸⁸ Pour comparer cette nouvelle approche à la théorie de la motivation, on peut dire que le désir n'est pas une attitude favorable envers un certain état de fait, mais qu'il s'agit plutôt d'une attitude envers un état de fait favorable. Si l'agent tient pour acquis qu'une chose est bonne et qu'une action donnée fera survenir cette chose, alors il en conclut que cette action est bonne. Le désir contient en lui-même la force justificatrice d'un jugement de valeur et fait partie intégrante du raisonnement derrière l'action. L'attitude motivationnelle de l'agent est donc intégrée dans sa raison d'agir. Velleman explique que selon les partisans de l'approche évaluative, puisque les désirs constituent des jugements de valeur en faveur de leur objet et qu'ils font partie du raisonnement menant à l'action, il en résulte que toutes les actions faites pour des raisons sont faites en vue du bien, puisqu'elles surviennent à la suite de jugements de valeur.

L'objection de Velleman contre cette théorie est la suivante : si les désirs sont des attitudes envers des propositions évaluatives et qu'ils contiennent en eux la force justificatrice d'un jugement de valeur, alors la capacité de désirer impliquerait la possession de concepts évaluatifs.⁸⁹ Il faudrait qu'un agent possède le concept de désirabilité, c'est-à-dire de ce qui vaut la peine d'être désiré, avant d'être en mesure de désirer quelque chose. Il faudrait donc qu'il ait une compréhension de ce qu'est le bien. Or, selon Velleman, il est clair que certains agents ne possèdent pas de tels concepts, bien qu'ils possèdent la capacité de désirer. Un jeune enfant, par exemple, peut fortement désirer avoir un jouet, mais il n'a pas la capacité mentale de comprendre en quoi il est désirable d'avoir ce jouet. Cela ne l'empêche pas de le désirer et d'être motivé à atteindre son but, qui est d'avoir le jouet. Il n'a simplement pas conscience que ce but vaut la peine d'être atteint, que le jouet vaut la peine d'être désiré. Le concept de désirabilité, selon Velleman, est quelque chose qui doit être enseigné à l'enfant. On lui apprend ce concept en disciplinant ses désirs existants.⁹⁰ Les désirs sont donc antérieurs, chez l'enfant, à la possession de concepts évaluatifs. Pour cette raison, l'approche évaluative est fautive

⁸⁸ *Ibid.*, p. 7.

⁸⁹ *Id.*

⁹⁰ *Id.*

parce qu'il est possible de désirer quelque chose sans avoir la conception que cette chose est bonne et vaut la peine d'être désirée.

Comment répondre à cet argument? Il faut tout d'abord se demander ce que signifie le fait de posséder des concepts évaluatifs. Il est indéniable que certaines capacités évaluatives sont requises pour qu'un agent soit capable d'agir en vue du bien. En effet, l'approche évaluative affirme que ces capacités jouent un rôle fondamental dans le processus menant à l'action intentionnelle. Par contre, la manière dont on définit ces capacités déterminera dans une large mesure les chances de survie de l'approche évaluative. Si le fait de posséder des concepts évaluatifs implique la capacité de penser de manière abstraite, de maîtriser des concepts philosophiques tels que le bien et la valeur et de les inclure dans une réflexion rationnelle précédant l'action, alors l'approche évaluative est vouée à l'échec. Personne ne peut sérieusement prétendre que ces éléments sont toujours présents et jouent un rôle chaque fois qu'un agent désire quelque chose ou accomplit une action intentionnellement. Nous avons mentionné dans le premier chapitre qu'il fallait éviter de proposer des théories aussi contraignantes, puisqu'elles sont facilement réfutables. Velleman, pour sa part, prend pour acquis que les capacités intellectuelles qu'un agent devrait posséder pour pouvoir évaluer qu'une chose est bonne sont suffisamment sophistiquées pour que certains agents, comme les jeunes enfants, soient incapables d'évaluer, bien qu'ils soient capables de désirer. Selon lui, le concept de désirabilité doit être enseigné aux enfants en disciplinant leurs désirs existants.⁹¹ Un partisan de l'approche évaluative comme Tenenbaum, pour sa part, penserait qu'il est impossible d'avoir la capacité de désirer sans avoir la capacité d'évaluer, puisque le désir est par nature un état dans lequel une chose nous apparaît comme bonne, donc comme valant la peine d'être désirée.

Un partisan de l'approche évaluative devra donc défendre l'idée que la capacité d'évaluer qu'une chose est bonne ne nécessite pas de facultés mentales avancées et que les agents les moins développés du point de vue cognitif, tels que les jeunes enfants et possiblement les animaux, possèdent cette capacité et s'en servent pour guider leur action. Trois

⁹¹ *Id.*

philosophes contemporains ont développé des idées allant en ce sens. Daniel Friedrich, dans « The Alluringness of Desire », ⁹² cherche à donner une explication non cognitiviste du fait que les choses que l'on désire semblent posséder un certain attrait à nos yeux. Il est impossible, selon lui, de désirer une chose tout en ayant une attitude entièrement négative, ou même entièrement neutre, vis-à-vis de cette chose. Pourtant, les explications cognitives de ce phénomène font face à la difficulté soulignée par Velleman : elles demandent un niveau de sophistication mentale que certains agents capables de désirer, tels que les jeunes enfants et les animaux, n'ont pas. De plus, une explication cognitiviste rend le cas des désirs récalcitrants problématique. Si désirer une chose implique d'avoir la croyance que cette chose est bonne, alors comment peut-on continuer de désirer quelque chose quand on a acquis la croyance que cette chose est mauvaise? Ceci semble contradictoire, mais dans les faits, les situations de ce genre sont courantes.

Une approche cognitiviste, affirme Friedrich, est incapable de résoudre ces problèmes. Il privilégie plutôt une explication non cognitiviste pour expliquer l'attrait caractéristique des désirs. Les états mentaux non cognitifs sont simplement des états affectifs qui nous font ressentir certains sentiments (*feel a certain way*⁹³). De tels états mentaux peuvent accompagner le contenu d'un état mental cognitif et lui donner un certain aspect positif ou négatif, indépendamment du contenu cognitif lui-même.⁹⁴ Ceci constitue, selon Friedrich, une évaluation non cognitive. Le désir implique donc une certaine attitude évaluative positive envers l'objet désiré. Cette attitude est une attraction non cognitive qui serait, comme l'illustre Friedrich, une sorte de « pouce levé » mental qui accompagne la représentation mentale de l'objet.⁹⁵ Cette explication a l'avantage de ne pas exiger que l'agent possède une très grande sophistication mentale pour être en mesure d'évaluer une chose comme étant bonne. Elle permet aussi d'expliquer comment nos désirs peuvent ne pas toujours s'aligner avec nos croyances concernant le bien.

⁹² Friedrich, 2012, p. 291.

⁹³ *Ibid.*, p. 297.

⁹⁴ *Ibid.*, pp. 297-298.

⁹⁵ *Ibid.*, p. 299.

Jessica Moss emprunte une approche différente dans son article « Aristotle's Non-Trivial, Non-Insane View That Everyone Always Desires Things Under the Guise of the Good », ⁹⁶ où elle propose une solution cognitiviste selon laquelle les sensations de plaisir sont en soi des perceptions cognitives de la valeur, et que par conséquent tous les désirs sont orientés vers le bien, par le biais de ce qui est plaisant. Il est important de souligner que Moss n'endosse pas personnellement cette théorie, mais qu'il s'agit plutôt d'une interprétation de la pensée d'Aristote qui était, selon elle, partisan d'une version radicale de l'approche évaluative. ⁹⁷ L'idée générale derrière la solution cognitiviste proposée par Moss est que lorsque l'on ressent du plaisir, cela revient à percevoir la cause de ce plaisir comme un bien. Ainsi, lorsqu'on désire ce qui est plaisant, on désire par le fait même ce qui est bien, parce que le plaisir est une forme d'appréhension du bien. Cette interprétation se base sur un passage d'Aristote dans *De Anima* où celui-ci affirme que lorsqu'une chose est plaisante ou douloureuse, alors l'âme, approuvant ou désapprouvant la sensation, cherche à poursuivre ce qui est plaisant et à fuir ce qui est douloureux. ⁹⁸ Ainsi, le plaisir et la douleur sont des sensations qui permettent à l'esprit humain de discriminer entre ce qui est bon et ce qui est mauvais, entre ce qui vaut la peine d'être recherché et ce qu'il vaut mieux éviter. Ceci fait du plaisir une sorte de cognition, un jugement de valeur. Si ressentir du plaisir est réellement une façon dont on peut avoir conscience de la valeur d'une chose, alors la capacité de désirer une chose parce qu'elle est bonne ne requiert pas une grande sophistication cognitive. ⁹⁹ Même les bébés et les animaux, qui ne possèdent pas de concepts ou de croyances à propos du bien, seraient en mesure d'agir en prenant le bien comme but.

Jennifer Hawkins propose un point de vue similaire dans « Desiring the Bad under the Guise of the Good ». ¹⁰⁰ Selon elle, les états mentaux que l'on peut qualifier de désirs sont également des états évaluatifs. Même les désirs non motivés qui ne sont pas justifiés par des raisons, mais qui sont tout simplement des états dans lesquels on se trouve, ont un aspect évaluatif. Ce ne sont pas uniquement des envies inintelligibles qui nous poussent à

⁹⁶ Moss, 2010.

⁹⁷ *Ibid.*, p. 2.

⁹⁸ *Ibid.*, p. 5.

⁹⁹ *Ibid.*, pp. 4;9.

¹⁰⁰ Hawkins, 2008, pp. 245-246.

faire quelque chose sans que l'on sache pourquoi. Au contraire, l'idée de poursuivre nos désirs semble toujours avoir un certain sens à nos yeux, même lorsque notre pensée rationnelle intervient pour juger que cela ne constitue pas une conduite appropriée. Les désirs non motivés ne sont certainement pas des croyances évaluatives, selon Hawkins, mais plutôt des perceptions ou des impressions évaluatives.¹⁰¹ En tant que tels, ils ne nécessitent pas de maîtriser le concept de bien. Tout comme la perception visuelle requiert certains concepts ou proto-concepts tels que la reconnaissance des objets dans l'espace, la capacité de désirer demande tout de même une certaine forme primitive de cognition.¹⁰² La capacité d'avoir des impressions évaluatives nécessite certains proto-concepts de bien, d'approprié, de raison, etc. Ces proto-concepts proviennent d'une expérience affective antérieure au développement des capacités conceptuelles matures. L'affect se définit, selon Hawkins, comme un certain type de réaction subjective, non sensorielle et continue par rapport au monde.¹⁰³ Les représentations affectives sont en mesure de présenter certaines choses comme étant bonnes, semblant correctes ou ayant un sens, d'une manière très primitive qui ne requiert pas de posséder des capacités conceptuelles développées. Certaines expériences affectives primaires seraient donc des impressions évaluatives. Le désir, par exemple, est une « bonne » impression évaluative puisqu'il présente son objet comme un bien. Hawkins soutient que même après que nos capacités conceptuelles soient pleinement développées, l'affect continue de faire partie de nos vies et de nous fournir certaines impressions de valeur.¹⁰⁴ Ces impressions ne sont pas assujetties à nos croyances, ce qui explique notamment la manière dont nos désirs peuvent parfois entrer en conflit avec ce que nous croyons être le bien.

Malgré leurs différences, Friedrich, Moss et Hawkins proposent des idées similaires, notamment celle selon laquelle la capacité d'évaluer et d'agir en fonction de cette évaluation ne nécessite pas des capacités cognitives aussi sophistiquées que le pense Velleman. Au contraire, c'est une capacité tout à fait primaire, intrinsèquement liée à la manière dont on perçoit et ressent le monde. Cette conception alternative peut-elle

¹⁰¹ *Ibid.*, pp. 247-248; 258-260.

¹⁰² *Ibid.*, pp. 254-256; Bermudez, 1998, ch. 3-4.

¹⁰³ Hawkins, 2008, p. 258.

¹⁰⁴ *Ibid.*, p. 262.

défendre l'approche évaluative contre l'argument de Velleman? Je pense que oui. La capacité d'appréhender des choses comme bonnes ou mauvaises joue un rôle tellement essentiel dans notre conduite qu'il est difficile de l'imaginer autrement que directement liée à la capacité d'agir intentionnellement. Même un bébé qui rejette une nourriture déplaisante et qui dévore avec enthousiasme une nourriture plaisante passe une forme de jugement évaluatif en agissant ainsi. Même un écureuil qui prend peur et qui s'enfuit lorsqu'un être humain s'approche de lui doit former un certain jugement qu'il vaut mieux pour lui de ne pas rester là. La façon la plus plausible d'expliquer les attitudes primitives d'attraction et de répulsion, omniprésentes chez les êtres vivants suffisamment développés, est de faire intervenir d'une quelconque manière la notion de valeur.

Prenons, par exemple, une sensation comme la soif. Si cette sensation n'était pas accompagnée d'une certaine impression évaluative, si elle n'était pas désagréable et ne nous plaçait pas dans un état d'inconfort qui nous pousse à vouloir revenir à un état plus agréable, alors nous traiterions la soif comme n'importe quelle autre sensation neutre. Nous ne nous en préoccuperions pas et nous ne serions pas motivés à l'étancher en buvant de l'eau. La soif est en soi une sensation non évaluative, mais elle est accompagnée d'une impression évaluative négative qui nous indique qu'il vaut mieux faire cesser cette sensation en étanchant notre soif. Sans cette évaluation, sans cette notion de valeur, il est difficile de concevoir que nous ferions quoi que ce soit de nécessaire à notre survie. Les êtres humains, comme les autres animaux, ne survivraient pas longtemps s'ils n'avaient pas la capacité de discriminer entre le bon et le mauvais. Cette capacité ne peut pas être uniquement à la portée d'êtres intellectuellement sophistiqués. Elle est cruciale et nécessaire à la survie de tous les animaux, et particulièrement de ceux dont l'intellect n'est pas suffisamment développé pour maîtriser des concepts abstraits. C'est pourquoi nous devons accepter, jusqu'à preuve du contraire, l'explication donnée par les partisans de l'approche évaluative selon laquelle la capacité d'évaluer et d'appréhender une chose comme étant bonne ou mauvaise est beaucoup plus primitive et universelle que le pense Velleman. Il n'est pas besoin de maîtriser le concept de bien, au sens où l'entend Velleman, pour pouvoir agir en accord avec l'approche évaluative.

2.3 Le problème de la croyance

Kieran Setiya développe deux arguments contre l'approche évaluative dans son livre *Reasons without Rationalism*¹⁰⁵ et dans un article intitulé « Sympathy for the Devil »¹⁰⁶. Comme on l'a vu précédemment, Setiya est d'avis que les contre-exemples ne peuvent pas, à eux seuls, réfuter définitivement l'approche évaluative puisqu'ils demeurent sujets à interprétation.¹⁰⁷ Il souhaite donc présenter des arguments fondés sur des principes abstraits, plutôt que sur des exemples concrets. Le premier argument de Setiya se fonde sur ce qu'il nomme le problème de la croyance. Avec cet argument, il cherche à démontrer que certaines contraintes doivent nécessairement s'appliquer sur une théorie de l'action intentionnelle plausible et que l'approche évaluative ne peut pas satisfaire ces contraintes, ce qui entame grandement la crédibilité de la théorie.¹⁰⁸

La contrainte que Setiya veut imposer à une théorie de l'action intentionnelle est la contrainte de la croyance. Il s'inspire de l'idée d'Elizabeth Anscombe à propos de la connaissance de nos propres actions. Cette idée, développée par Anscombe, est que nous avons une connaissance sans observation (*knowledge without observation*) de nos actions intentionnelles.¹⁰⁹ Nous connaissons nos propres actions sans avoir besoin de les observer. Lorsque nous agissons intentionnellement, nous avons la croyance que nous faisons X et cette croyance s'acquiert de manière non inférentielle, c'est-à-dire qu'elle survient spontanément sans être le résultat d'une quelconque interaction avec le monde. Bien que certaines exceptions soulevées par Donald Davidson¹¹⁰ exigent que cette idée soit qualifiée, Setiya admet qu'il y a au moins une part de vrai dans l'idée d'Anscombe. Lorsque nous agissons intentionnellement, nous devons avoir la croyance que nous sommes en train de faire X. C'est une vérité nécessaire en théorie de l'action.

¹⁰⁵ Setiya, 2007.

¹⁰⁶ Setiya, 2010.

¹⁰⁷ Setiya, 2007, p. 22; Setiya, 2010, p. 83.

¹⁰⁸ Setiya, 2007, p. 22.

¹⁰⁹ Anscombe, 1963, pp. 13-15; Setiya, 2007, pp. 24-25; Tenenbaum, 2012, p. 91.

¹¹⁰ Davidson, 1980, p. 50; 91-94.

Puisque c'est une vérité nécessaire que nos actions intentionnelles s'accompagnent de la croyance que nous agissons ainsi, affirme Setiya, alors une théorie de l'action intentionnelle crédible doit pouvoir l'expliquer et en rendre compte. Une théorie acceptable devra donc satisfaire au critère de la croyance. Selon Setiya, l'approche évaluative ne permet pas d'expliquer ce critère, mais il existe des théories plausibles qui satisfont le critère de la croyance sans avoir besoin de postuler l'idée que nous agissons pour des raisons qui nous semblent bonnes. L'approche évaluative en tant que théorie de l'action intentionnelle est donc superflue et probablement fausse.

Pourquoi l'approche évaluative ne satisfait-elle pas le critère de la croyance? Selon Setiya, la manière dont la « Guise of the Good Theory » explique l'action intentionnelle consiste à dire que lorsqu'un agent agit pour une raison, cela signifie qu'il *prend* quelque chose comme *sa* raison d'agir. (*Taking something as one's reason*) De plus, selon la « Guise of the Good Theory », prendre quelque chose comme une raison d'agir signifie que l'agent a une certaine attitude de croyance-désir (*Desire-like belief*) envers sa raison qui la représente comme une bonne raison d'agir.¹¹¹ Le défaut d'une telle explication, pense Setiya, est qu'elle n'explique pas suffisamment bien la croyance que nous avons concernant nos propres actions. Le fait que je crois qu'une raison P est une bonne raison de faire une action X n'implique pas que je crois être en train de faire X. La croyance que P est une bonne raison de faire X n'est pas intrinsèquement liée avec la croyance que je fais X. Il est tout à fait possible d'avoir l'une sans l'autre, donc de croire qu'il existe une bonne raison de faire quelque chose sans toutefois croire que je suis en train de faire cette chose. L'approche évaluative ne donne pas une explication suffisante de la croyance, elle la rend mystérieuse.¹¹²

Setiya propose une théorie causale de l'action faite pour une raison qui satisfait le critère de la croyance sans pour autant impliquer la vérité de l'approche évaluative.¹¹³ Selon lui, tenir quelque chose pour une raison d'agir signifie avoir une attitude de croyance-désir

¹¹¹ Setiya, 2007, pp. 39-41. « There is an obvious suggestion here, for those who defend the guise of the good: to take something as one's reason is to have the desire-like belief that it is a *good* reason on which to act. »

¹¹² *Ibid.*, p. 41.

¹¹³ *Ibid.*, pp. 56-59.

que je fais présentement l'action X à cause d'une croyance envers une raison P.¹¹⁴ Une telle explication résout le problème de la croyance, puisque si l'on croit faire X à cause de la raison que P, il s'ensuit nécessairement que l'on croit faire X.¹¹⁵ Ces explications peuvent sembler très abstraites, mais ce que Setiya cherche à faire ici est donner une définition purement explicative du concept de raison d'agir, tandis que les partisans de l'approche évaluative ont tendance à donner une définition normative qui justifie l'action aux yeux de l'agent. Là se trouve l'erreur, puisqu'une définition normative rend mystérieuse la croyance que nous accomplissons nos actions, tandis qu'une définition purement explicative de nos raisons d'agir permet de rendre compte adéquatement de cette croyance. En conclusion, Setiya affirme que puisque l'approche évaluative échoue à satisfaire le critère de la croyance et qu'une autre théorie est en mesure de le faire sans invoquer un élément normatif dans l'explication de nos raisons d'agir, cela constitue un argument contre l'approche évaluative en tant que théorie de l'action intentionnelle.

Il est important de souligner que Setiya ne prend pas clairement position sur la question de l'épistémologie de la croyance de nos actions. Il n'endosse pas complètement l'idée d'Anscombe selon laquelle la connaissance de nos actions s'acquiert sans observation et qu'elle est non inférentielle. Dans le passage qui suit, cependant, c'est à l'idée que cette croyance ou cette connaissance est acquise sans observation que nous devons répondre, puisque c'est le seul type de connaissance qu'une théorie de l'action intentionnelle doit être en mesure d'expliquer. En effet, si une connaissance s'acquiert par l'observation, alors l'observation constitue une explication satisfaisante de cette connaissance et la théorie n'a pas besoin de faire un effort supplémentaire pour l'expliquer. C'est seulement si on pense que la connaissance de nos actions s'acquiert sans observation que l'on peut exiger qu'une théorie de l'action intentionnelle explique comment elle s'acquiert. L'argument de Setiya doit donc reposer, malgré sa position agnostique sur la question, sur la notion que nous avons une connaissance sans observation de nos actions intentionnelles.

¹¹⁴ *Ibid.*, p. 46. « To take *p* as one's reason for doing ϕ is to have the desire-like belief that one is hereby doing ϕ because of the belief that *p*. »

¹¹⁵ *Ibid.*, p. 42.

L'argument de Setiya consiste à stipuler qu'une contrainte doit peser sur toute théorie de l'action intentionnelle plausible puis à démontrer que l'approche évaluative ne satisfait pas à cette contrainte. Il existe deux stratégies possibles pour répondre à cet argument. La première stratégie, utilisée par Sergio Tenenbaum, consiste à répondre que l'approche évaluative satisfait bel et bien la contrainte de la croyance et qu'elle est en mesure d'expliquer la connaissance sans observation de nos propres actions. Dans un article intitulé « *Knowing the Good and Knowing What One is Doing* », Tenenbaum s'affaire à prouver que l'approche évaluative implique ce qu'il appelle le réflexivisme, l'idée qu'un agent agit intentionnellement s'il a une connaissance sans observation du fait qu'il agisse ainsi.¹¹⁶ Tenenbaum explique que l'approche évaluative, qu'il nomme la thèse scolastique, implique la vérité du réflexivisme, puisque pour pouvoir être des agents qui fondent leurs actions sur une connaissance pratique du bien, il est nécessaire de connaître non seulement nos actions, mais également les raisons derrière celles-ci. La connaissance de nos actions intentionnelles est un élément nécessaire pour pouvoir guider notre conduite en vue de ce que nous percevons comme étant le bien. Il est difficile d'imaginer qu'un agent puisse être guidé par une conception complexe du bien tout en n'étant pas au courant de ce qu'il est en train de faire.¹¹⁷ Il ne serait pas raisonnable de penser que des agents suffisamment complexes pour pouvoir agir en vue du bien seraient incapables de connaître leurs propres actions. Cela serait absurde. Il semble que la connaissance de nos actions est un élément nécessaire à la capacité d'orienter nos actions en vue du bien. Nous ne pourrions pas agir de la manière décrite par la théorie scolastique si nous n'avions pas cette connaissance. Ainsi, puisque la connaissance de nos actions est nécessaire pour pouvoir agir en vue du bien, et que la théorie scolastique affirme que nous agissons en vue du bien, il est nécessaire qu'elle implique aussi la vérité du réflexivisme. Tenenbaum considère donc que la théorie scolastique, ou l'approche évaluative, satisfait le critère de la croyance imposé par Setiya.

La deuxième stratégie possible pour répondre à l'argument de Setiya consiste à montrer que la contrainte imposée par Setiya n'est pas justifiée et que l'approche évaluative n'a

¹¹⁶ Tenenbaum, 2012, pp. 93-95.

¹¹⁷ *Ibid.*, p. 108.

pas besoin de s'y conformer. Dans le reste de cette section, je tenterai de démontrer que le réflexivisme est faux et qu'il n'y a pas, en fait, de connaissance de nos actions sans observation. Si j'ai raison, alors l'argument de la croyance échoue. Il n'est pas une contrainte pour l'approche évaluative ni pour n'importe quelle autre théorie de l'action.

L'argument qui suit cherche à prouver que la connaissance de nos actions s'acquiert entièrement ou en majeure partie par l'observation et que l'idée que nous puissions avoir une connaissance de nos actions sans observation est fautive. Cet argument est basé sur des exemples et sur les conclusions que l'on peut en tirer. Imaginons que j'accomplisse une action tout à fait simple et triviale, par exemple, que j'écrive à la main la lettre « A » sur une feuille de papier. Comment puis-je connaître mon action? Tout simplement en observant ce que je suis en train de faire. Je vois la lettre se dessiner sous mes yeux et je vois que c'est ma main qui est responsable de cette action. Je connais donc parfaitement bien mon action : je sais que je suis en train d'écrire la lettre « A » sur une feuille de papier. Maintenant, imaginons que j'essaie de faire la même action, cette fois en gardant les yeux fermés. La connaissance de ce que je suis en train de faire est grandement diminuée. Je sais que je tiens un objet dans mes mains, dont j'ai le souvenir qu'il s'agit d'un crayon, et que je le bouge de manière à dessiner une forme approximative sur une surface plane. Déjà, la connaissance de ce que je fais est beaucoup moins détaillée. La seule chose dont j'ai vraiment une connaissance certaine est que je tiens quelque chose dans mes mains et que je bouge ma main d'une certaine manière. Je peux tout à fait former l'intention dans mon esprit d'écrire la lettre « A » sur une feuille de papier, mais tandis que je tente d'accomplir cette action, ma connaissance de ce que je suis effectivement en train de faire est bien plus limitée lorsque mes yeux sont fermés que lorsqu'ils sont ouverts. Il ne fait aucun doute, alors, que l'observation de ce que je suis en train de faire contribue à la connaissance que j'ai de mon action. Si on utilise un exemple d'action plus complexe que le simple fait d'écrire la lettre « A » sur une feuille de papier, la vérité de cette affirmation est encore plus évidente.

Mais alors, ne pourrait-on pas répondre qu'en fait, la seule véritable connaissance de nos actions est cette connaissance primaire qui se limite aux seuls mouvements du corps, et

que toute compréhension plus poussée de ce que nous faisons est superposée sur cette connaissance primaire de nos mouvements, qui est une connaissance sans observation? Cette idée fait face à deux problèmes. Le premier, c'est que même ceux qui croient à l'existence d'une connaissance sans observation seraient sans doute réticents à admettre que la connaissance de nos actions se limite effectivement aux simples mouvements de notre corps. Ils devraient soit admettre que la connaissance de nos propres actions n'inclut pas certaines connaissances plus complexes qui dépassent le simple mouvement du corps, comme faire chauffer le four à 400 degrés ou s'arrêter à une intersection pour laisser des piétons traverser la rue, ce qui est absurde; ou encore, affirmer que cette connaissance plus complexe s'acquiert aussi sans observation, ce qui est tout aussi absurde.

Le second problème est que même la connaissance des mouvements de notre corps s'acquiert, au moins en partie, par l'observation. Il faut garder à l'esprit que le terme « observation » employé dans ce contexte ne se limite pas uniquement à l'observation visuelle. Si une connaissance s'acquiert par « l'observation », elle peut s'acquérir aussi bien par la vue, par l'ouïe, par le toucher ou par n'importe quel autre sens. Si je marche et que mes yeux sont fermés, je peux savoir que je marche par la sensation de mes pieds foulant le sol, ou encore par le bruit de mes pas. Par contre, si j'imagine que ces stimuli sensoriels ne me sont plus accessibles, je vois mal comment je pourrais savoir que je suis en train de marcher même si je forme cette intention et que je crois mettre mon corps en mouvement. Je pourrais bel et bien être en train de marcher, mais sans aucune confirmation sensorielle de ce que je suis en train de faire, je n'aurais pas connaissance de ce fait. La connaissance de mon action s'acquiert, littéralement, par l'observation fournie par les sens, dans ce cas-ci la vue, l'ouïe et le toucher. Une connaissance de l'action acquise strictement sans observation apparaît tout simplement inconcevable.

Nous avons donc de bonnes raisons de penser que la connaissance de nos actions s'acquiert, totalement ou en grande partie, par l'observation de ce que nous faisons, ce qui signifie que cette connaissance survient sur l'action grâce à l'observation et qu'elle est donc une conséquence de l'action. Par conséquent, une théorie de l'action

intentionnelle n'a pas besoin d'expliquer la connaissance de nos actions. Peu importe la manière dont une théorie explique l'action intentionnelle, la connaissance de nos actions s'acquiert toujours de la même manière, c'est-à-dire indirectement, par l'observation de ce que nous faisons. Setiya croit attaquer l'approche évaluative en disant que s'il existe une connexion entre cette théorie et la connaissance nos actions, cette connexion est indirecte,¹¹⁸ mais en réalité c'est justement le type de connexion qui doit exister entre la genèse de nos actions et la connaissance que nous en avons.

Nous avons exploré deux réponses possibles à l'argument de la croyance. Tenenbaum, pour sa part, cherche à démontrer par son argument que l'approche évaluative répond bel et bien à la contrainte proposée par Setiya. Pour ma part, je trouve plus plausible d'argumenter contre l'idée que la connaissance de nos actions s'acquiert sans observation. Indépendamment du point de vue que l'on adopte sur le sujet de la connaissance de nos actions, l'approche évaluative possède les ressources nécessaires pour résister à l'argument fondé sur le problème de la croyance.

2.4 Les raisons non normatives

La dernière objection contre l'approche évaluative à considérer est un autre argument présenté par Kieran Setiya dans *Reasons without Rationalism* et « Sympathy for the Devil ». ¹¹⁹ Il est possible, selon lui, pour un agent d'agir pour une raison et de reconnaître qu'il agit pour une raison, sans toutefois penser que cette raison justifie l'action d'une quelconque manière ou qu'elle présente l'action sous un jour favorable.¹²⁰ Si Setiya a raison, cela signifie que l'approche évaluative est fautive, parce qu'il serait alors possible d'agir non seulement intentionnellement, mais aussi, jusqu'à un certain point, rationnellement sans faire intervenir un jugement de valeur ou une quelconque évaluation positive de l'action accomplie.

¹¹⁸ Setiya, 2010, p. 89.

¹¹⁹ Setiya, 2007; Setiya, 2010.

¹²⁰ Setiya, 2007, p. 60 « in general, we can say what is involved in acting for a reason without appeal to normative or evaluative thoughts. I can act because I take something as my reason for acting, and thus act for that reason, without taking the reason to be good. » ; Setiya, 2010, p. 90.

Setiya commence par définir deux types de raisons et donne une explication de ce en quoi agir pour une raison consiste. Il fait la distinction entre raisons normatives et raisons psychologiques. Une raison normative est un fait qui justifie objectivement une action. Une raison psychologique est une raison donnée par l'agent pour expliquer sa motivation à poser une action X.¹²¹ Il n'y a pas de lien nécessaire entre ces deux types de raisons et dans le présent débat, seules les raisons psychologiques nous intéressent. Une raison psychologique est ce qui constitue, selon Setiya, une réponse valide au « Pourquoi? » d'Anscombe.¹²² De plus, il définit l'attitude d'agir pour une raison en termes de « prendre P comme raison de faire X » (*Taking as one's reason*). Ainsi, lorsque je prends P comme raison d'agir, P est aussi la réponse que je donnerais au « Pourquoi? » d'Anscombe.¹²³

Avec ces concepts en main, Setiya affirme que les raisons psychologiques sont, en règle générale, des raisons purement explicatives, pas des raisons justificatives. Ainsi, si un agent donne une raison psychologique sous la forme de l'affirmation : « Je fais X parce que P », il n'est pas nécessaire, comme l'affirme l'approche évaluative, que P présente l'action X favorablement aux yeux de l'agent. L'agent peut croire qu'il prend P comme raison de faire X sans pour autant penser que P constitue une quelconque justification en faveur de l'action X. La réponse fournie par l'agent est strictement explicative et ne fait pas intervenir de jugement normatif. L'argument de Setiya consiste en quatre affirmations :¹²⁴

- 1) Il est suffisant pour répondre à la question « Pourquoi? » que l'agent ait une croyance telle que « Je fais X parce que P ».
- 2) Le fait que je fais X parce que P (raison psychologique) est cohérent avec le fait que P n'est pas une raison de faire X (raison normative).
- 3) Si une proposition est cohérente avec la négation d'une autre, il est possible de croire la première sans croire la seconde.

¹²¹ Setiya, 2007, pp. 28-29; Setiya, 2010, p. 87.

¹²² Setiya, 2007, p. 41; Setiya, 2010, p. 89.

¹²³ Setiya, 2010, p. 90. « It is sufficient to answer the question « Why? » that one has a belief of the form « I am doing ϕ because P » ».

¹²⁴ *Ibid.*, p. 90.

4) Il est possible de croire que je fais X parce que P (raison psychologique) et répondre à la question « Pourquoi? », sans toutefois croire que P est une raison de faire X (raison normative).

Cet argument sert à démontrer que nous pouvons croire que l'on fait X parce que P sans toutefois s'engager par rapport au statut de P en tant que raison normative de faire X. La conséquence de l'argument est que nous pouvons reconnaître que nous agissons à cause d'une raison et nier que cette raison justifie de quelque manière que ce soit notre action, même à nos propres yeux. Setiya illustre ce point à l'aide d'un exemple. Admettons qu'un homme sorte en courant d'une maison en feu. À ce moment-là, on peut lui demander pourquoi il sort de la maison. Celui-ci répond que c'est parce que la maison est en feu. Il serait alors possible de rétorquer que sa famille dort présentement à l'étage, et que le fait que la maison soit en feu constitue une raison de secourir ses proches plutôt qu'une raison de sortir. L'homme répondrait alors qu'en effet, le fait que la maison soit en feu ne constitue pas une raison de sortir, mais que c'est tout de même la raison pour laquelle il sort.¹²⁵

À la lumière de cet exemple, Setiya affirme que l'approche évaluative est fautive puisque rien ne permet de démontrer la nécessité d'un élément évaluatif dans le raisonnement menant à l'action. Setiya pousse son raisonnement encore plus loin, puisqu'il conclut qu'il est non seulement possible d'agir pour une raison qui ne nous paraît pas bonne, mais que c'est ce qui se produit de manière générale quand on agit. On peut généralement expliquer ce en quoi consiste le fait d'agir pour une raison sans faire référence à des éléments normatifs ou évaluatifs. Sur ce dernier point, Setiya s'oppose directement à Joseph Raz, puisqu'il accuse celui-ci de faire une pétition de principe en affirmant que des raisons explicatives doivent nécessairement être comprises en termes normatifs.¹²⁶

L'argument de Setiya en faveur des raisons explicatives fait face à trois problèmes. Le premier est qu'il y a une faute dans la structure logique de l'argument. Il ne fait aucun

¹²⁵ *Ibid.*, p. 90.

¹²⁶ Setiya, 2007, p. 60.

doute que la deuxième affirmation de Setiya est correcte, et que le fait qu'un agent fasse X pour la raison que P est cohérent avec le fait que P n'est pas une raison normative de faire X. Le problème survient lors du passage de la deuxième vers les troisième et quatrième affirmations, puisque bien qu'il soit possible qu'un agent croie quelque chose et qu'il ait objectivement tort de croire cela, cela ne veut pas dire qu'un seul et même agent puisse croire en même temps deux affirmations contradictoires. En effet, il est possible que les propositions « un agent croit que X est vrai » et « X est faux » soient toutes les deux vraies, mais un seul et même agent ne peut pas à la fois croire « X est vrai » et « X est faux » sans contradiction. L'argument de Setiya semble faire cette erreur puisqu'il prend, à la deuxième affirmation, le fait objectif que P n'est pas une raison normative de faire X, et le transforme en une croyance dans les affirmations trois et quatre. Ainsi, les deux affirmations cohérentes du point deux de l'argument, soit « l'agent croit "je fais X parce que P" (au sens où P est sa raison de faire X) » et « P n'est pas une raison de faire X », deviennent deux affirmations potentiellement incohérentes aux points trois et quatre, soient « l'agent croit "je fais X parce que P" (au sens où P est sa raison de faire X) » et « l'agent croit "P n'est pas une raison de faire X" ».

Il n'est pas certain, cependant, que ces deux affirmations doivent nécessairement être incohérentes. Après tout, un agent pourrait penser que le concept de raison normative est suffisamment différent du concept de raison psychologique pour être en mesure de croire que P est une raison psychologique de faire X, mais pas une raison normative. C'est effectivement la position qu'adopte Setiya, puisqu'il est d'avis que les agents n'essaient pas nécessairement d'aspirer à un certain idéal de rationalité lorsqu'ils agissent intentionnellement.¹²⁷ Cependant, ce point n'est pas démontré par l'argument. Pour que l'argument soit valide, il faudrait accepter en tant que prémisse l'idée que l'agent lui-même puisse ne pas concevoir ses propres raisons d'agir comme des raisons normatives, mais ceci est loin d'être évident, et c'est justement ce que l'argument cherche à prouver. Setiya se rend donc lui-même coupable d'une pétition de principe, car la conclusion de son argument, qu'un agent peut croire que P est sa raison psychologique de faire X sans croire que P est une raison normative d'agir, dépend de la prémisse que l'agent pourrait

¹²⁷ *Ibid.*, pp. 66-67; Setiya, 2010, pp. 93-101

penser que ses propres raisons d'agir ne sont pas des raisons normatives, ce qui est justement l'hypothèse que Setiya tente de démontrer.

Le deuxième problème auquel Setiya fait face n'est pas un problème dans l'argument lui-même, mais plutôt dans la théorie que Setiya endosse. Accepter la théorie de Setiya revient à admettre que les raisons psychologiques sont, dans une large mesure, mystérieuses et inexplicables. Si on admet, comme Setiya, que l'on choisit les raisons pour lesquelles on agit¹²⁸ et que l'on peut généralement expliquer ces raisons sans faire référence à un aspect évaluatif, alors sur quelle base ce choix se fait-il? Quand on fait un choix, on choisit selon certains critères, et qui dit critères de sélection dit évaluation. Comment serait-il possible de faire un choix qui ne se base sur rien du tout? Comment choisir si on n'a pas le moindre critère en fonction duquel décider? Setiya dit que nous choisissons nos raisons d'agir, mais sa théorie ne fournit pas les outils qui nous permettraient de comprendre et d'expliquer ces choix. Sa théorie transforme les raisons psychologiques d'agir en états mystérieux qui surviennent sans que l'on sache pourquoi. La conséquence nécessaire d'une théorie qui élimine tout aspect évaluatif du processus de décision humain est une incapacité à comprendre de manière plus approfondie pourquoi nos raisons d'agir sont effectivement des raisons d'agir.

Prenons, par exemple, le scénario de la maison en feu développé par Setiya afin d'étayer son argument. Dans cet exemple, l'agent tient apparemment le fait que la maison soit en feu pour sa raison de sortir, même s'il ne pense d'aucune manière que c'est une bonne chose de sortir parce que la maison est en feu. Ainsi, sans faire intervenir une quelconque évaluation dans son processus de décision, il décide que le fait que la maison soit en feu est la raison pour laquelle il sort. Comment arrive-t-il à cette conclusion? Pourquoi le fait que la maison est en feu serait-il spécifiquement la raison prise par l'agent pour sortir? Pourquoi ne prendrait-il pas cela comme une raison de jouer aux cartes? Sans une évaluation de la part de l'agent, il n'y a absolument aucune manière pour lui de comprendre que le fait que la maison soit en feu est davantage une raison de s'enfuir qu'une raison de jouer aux cartes. La théorie de Setiya rend la décision de l'agent

¹²⁸ Setiya, 2007, pp. 39-40.

mystérieuse. Elle fournit peut-être une réponse au « Pourquoi? » d'Anscombe, mais elle n'explique pas le pourquoi de la raison elle-même. Cette théorie semble rendre impossible une explication des raisons d'agir plus poussée que la simple affirmation : « l'agent prend P comme raison de faire X. »

Pourtant, il doit être possible d'avoir une compréhension plus approfondie des raisons psychologiques d'agir. On peut très bien comprendre en quoi le fait que la maison soit en feu peut constituer davantage, aux yeux de l'agent, une raison de sortir que de jouer aux cartes. Cette compréhension repose sur l'évaluation, de la part de l'agent, qu'il est préférable de ne pas brûler vif. Pour pousser la compréhension de nos raisons d'agir plus loin que la simple affirmation que nous faisons X à cause de P, il est nécessaire de faire intervenir des éléments évaluatifs. Sans ces éléments, la compréhension du pourquoi de nos actions demeure extrêmement limitée. La théorie de Setiya ne présente qu'un niveau très superficiel de compréhension de nos actions et de nos raisons d'agir, et nie l'importance d'un élément pertinent à cette compréhension, soit l'évaluation faite par l'agent. Accepter la théorie de Setiya, c'est se priver d'une compréhension plus riche et plus complète de nos raisons d'agir.

Le troisième problème de l'argument de Setiya concerne l'idée plus générale que nous agissons pour des raisons. Bien qu'il soit tout à fait pertinent, pour des raisons pratiques, de s'exprimer en des termes tels que « Je fais X parce que P », il ne faut toutefois pas oublier qu'il s'agit d'une simplification de la réalité. On n'agit pas toujours pour une raison spécifique, mais plutôt en se fondant sur une vue d'ensemble de la situation. La compréhension de certains faits, que l'on nomme des raisons, peut contribuer, à un degré plus ou moins élevé, à notre décision d'agir d'une certaine manière. Mais il est erroné d'affirmer, comme le fait Setiya, que je prends le fait qu'il fasse beau dehors comme ma raison d'aller prendre une marche, que je prends une marche *parce qu'il* fait beau dehors.¹²⁹ D'autres facteurs contribuent à motiver cette décision, comme le fait que je n'aie rien de plus important à faire, que je me sente suffisamment en forme pour aller marcher, que j'ai envie de me dégourdir les jambes, etc. Tous ces faits contribuent à

¹²⁹ *Ibid.*, pp. 41-42.

rendre l'idée d'aller prendre une marche plus attrayante. La météo n'est pas le seul facteur qui me motive à aller marcher. Tout au plus, c'est un facteur décisif parmi une foule d'autres facteurs de moindre importance. Ce n'est pas parce qu'il fait beau dehors que je vais marcher. Même s'il fait beau dehors, si j'ai quelque chose de plus important à faire, ou si je n'ai pas envie de prendre une marche, alors je n'irai pas. Lorsque je suis motivé à aller marcher, c'est parce que plusieurs facteurs sont présents et contribuent à rendre l'idée d'aller marcher suffisamment attrayante pour me convaincre d'agir. La complexité inhérente à la notion de raisons psychologiques d'agir peut parfois être laissée de côté pour des raisons pratiques afin de ne pas compliquer la discussion outre mesure, mais il est tout de même nécessaire d'en tenir compte lorsqu'on souhaite formuler une explication théorique de ce que c'est d'agir pour des raisons.

Dès lors que nous acceptons que les motivations derrière nos actions soient plus complexes que « je fais X pour la raison que P », l'approche évaluative s'en sort beaucoup mieux que la théorie causale de Setiya. Plus notre compréhension de l'action intentionnelle se complexifie, plus la théorie de Setiya semble vouée à l'échec. Si cette théorie est incapable d'expliquer suffisamment bien pourquoi un agent prend P comme raison de faire X plutôt que comme raison de faire Y, comment peut-elle expliquer en quoi les facteurs Q, R, S et T contribuent à la décision de faire X, tout en niant que l'évaluation ait un rôle à jouer dans le processus de décision? Lorsque le tableau se complexifie, la théorie causale perd tout contact avec la réalité. Elle n'a tout simplement pas les outils nécessaires pour expliquer les motivations d'un agent intellectuellement sophistiqué opérant dans un monde dynamique et complexe.

Pour l'approche évaluative, l'effet contraire se produit. Lorsque les facteurs motivant une décision particulière sont nombreux et complexes, la nécessité d'un élément évaluatif capable de guider l'action prend tout son sens. Les capacités évaluatives nous permettent d'analyser une situation, de comparer les différents facteurs selon leur importance respective et de discriminer entre ce qui est bon et mauvais, utile et inutile, valable et sans valeur. Nous avons, à chaque instant, une compréhension dynamique du monde à l'intérieur de laquelle certaines actions possibles nous apparaissent bonnes ou mauvaises.

Cette évaluation constante est à la base de toutes nos décisions, que nous en ayons conscience ou pas. Il devient progressivement plus difficile de concevoir que l'évaluation puisse ne pas jouer de rôle dans nos prises de décisions à mesure que l'on adopte une vision plus complexe et détaillée de ce que c'est que d'agir intentionnellement. L'évaluation est intrinsèquement liée à l'action. Elle est l'outil qui nous permet de prendre des décisions, et donc d'agir. Ainsi, agir intentionnellement sans faire intervenir un quelconque élément évaluatif semble inconcevable.

Nous avons vu que l'argument de Setiya en faveur de l'existence de raisons non normatives est fort probablement basé sur un sophisme, puisqu'il prend comme prémisse implicite l'idée qu'un agent pourrait ne pas voir ses raisons psychologiques d'agir comme des raisons au sens normatif, ce qui est justement ce que Setiya doit prouver. De plus, au-delà de l'argument lui-même, nous avons vu que la théorie causale de Setiya fait face à de sérieux problèmes lorsqu'elle doit expliquer pourquoi un agent tient un fait particulier pour une raison d'accomplir une certaine action, et surtout comment un agent peut faire un tel choix sans une quelconque évaluation. Finalement, nous avons vu que les problèmes inhérents à la théorie proposée par Setiya deviennent encore plus importants à mesure que nous adoptons une vision plus complexe et plus réaliste de ce que c'est que d'agir pour des raisons et de prendre des décisions. L'approche évaluative, pour sa part, prend tout son sens lorsqu'il s'agit d'expliquer des décisions complexes, fondées sur une pluralité de facteurs. Ainsi, la tentative de Setiya de discréditer l'approche évaluative en proposant une théorie plus fonctionnelle qui laisserait de côté tout élément évaluatif dans l'explication de nos actions intentionnelles semble avoir échoué. L'approche évaluative demeure une meilleure théorie pour expliquer l'action intentionnelle que la théorie causale de Setiya.

CHAPITRE III

VERSION OPTIMALE DE L'APPROCHE ÉVALUATIVE

3.1 Pour une théorie sans faille

Dans le chapitre précédent, nous avons présenté et réfuté les objections adressées à l'approche évaluative. En l'absence d'autres objections, l'approche évaluative émerge comme la meilleure des théories de la motivation existantes. Cependant, nous devons encore nous pencher sur les détails d'une version optimale de l'approche évaluative. Nous avons vu, dans le premier chapitre, qu'il existe des différences notables entre les versions de l'approche évaluative défendues par Raz et par Tenenbaum. Au-delà de ces deux versions, plusieurs variantes sont également possibles. Dans ce chapitre, nous tenterons de découvrir quelle version de l'approche évaluative est la plus plausible.

Il faut tout d'abord se demander si l'approche évaluative doit admettre des exceptions au principe que les actions intentionnelles sont motivées par une évaluation. Si oui, ce sera une version modérée de la théorie. Sinon, ce sera une version radicale, dite « sans faille ». Dans cette section, nous défendrons l'idée que l'approche évaluative doit être sans faille et que toutes les actions intentionnelles sont accomplies *sub specie boni*, c'est-à-dire qu'elles sont motivées par une évaluation. Trois arguments seront présentés pour défendre cette affirmation. Ces arguments dérivent en partie du fait que l'on tienne pour vraie l'idée générale derrière l'approche évaluative, et cherchent à construire sur de telles bases. Le premier argument est que toute action intentionnelle implique un choix et que tout choix implique une évaluation. Le deuxième est qu'une explication intuitivement satisfaisante de l'action intentionnelle doit toujours faire référence à un bien appréhendé par l'agent. Le troisième est que l'existence d'un processus évaluatif primaire est nécessaire pour expliquer les actions faites sans réfléchir et que l'existence d'un tel processus rend possible l'explication de toutes les actions intentionnelles par l'approche évaluative. Ensemble, ces arguments démontrent qu'une version sans faille de l'approche évaluative est logique, intuitive et également compatible avec les actions triviales et irréfléchies que l'on pourrait être tenté d'exclure de la théorie.

Une action intentionnelle est une action résultant d'un choix, donc d'une décision d'agir d'une certaine manière. Pour agir intentionnellement, il faut vouloir accomplir l'action au moment où on l'accomplit. Ainsi, toute action intentionnelle implique un choix. Comme nous l'avons vu dans la section 2.4, il est difficile de concevoir qu'un choix puisse être fait sur la base de rien. Quand on choisit X plutôt qu'Y, on peut toujours se demander pourquoi on fait ce choix. Inévitablement, la réponse à cette question pointe vers une certaine considération qui fait pencher la balance davantage du côté de X que de celui d'Y. Pour être en mesure de choisir entre deux ou plusieurs options, un agent doit nécessairement se fonder sur une quelconque considération favorisant une option plutôt qu'une autre. Il doit y avoir quelque chose, dans l'esprit de l'agent, qui compte davantage en faveur de l'option X que de l'option Y, et on peut dire que ce quelque chose est la raison pour laquelle l'agent choisit X plutôt qu'Y. Pour nous convaincre de la vérité de cette affirmation, nous pouvons examiner les cas où nous avons de la difficulté à choisir. C'est souvent lorsqu'aucune option ne nous apparaît préférable à une autre que nous avons le plus de difficulté à faire un choix. Sans une idée claire de ce qui constitue la meilleure option, nous sommes parfois tout à fait incapables de choisir. L'agent doit donc être convaincu, d'une manière ou d'une autre, que X vaut mieux qu'Y pour être en mesure de choisir X, et cela constitue la base sur laquelle un choix peut se faire.

Ce qui constitue la base d'un choix est nécessairement une forme d'évaluation. Faire un choix, c'est discriminer entre plusieurs possibilités afin d'en sélectionner une et de rejeter les autres. Peu importe les considérations qui motivent l'agent à choisir une option plutôt qu'une autre, faire un choix revient finalement à affirmer que l'option choisie est meilleure que l'option rejetée. Ce principe s'applique également dans les cas où l'agent reconnaît volontiers qu'il choisit une option qui n'est pas idéale ou qui constitue le moindre mal. En effet, même si l'agent accepte de faire un compromis et ne choisit pas l'option qui serait la meilleure pour lui, à cause de considérations diverses, il est néanmoins porté à faire le choix qu'il considère comme optimal, toutes choses considérées. Il accorde de la valeur à différentes considérations, et ces considérations peuvent le convaincre qu'il est mieux de ne pas choisir uniquement en fonction de ses

préférences personnelles. Malgré tout, lorsque le jugement de l'agent s'arrête sur une option particulière, il faut admettre que cette option représentait la meilleure possible aux yeux de cet agent au moment où il a pris la décision, et que c'est pour cette raison qu'il l'a effectivement choisie. La notion de choix est donc logiquement liée, en fin de compte, à la notion d'évaluation. Même dans les cas où l'on choisit sur la base de considérations triviales qui n'ont aucun rapport à nos yeux avec le concept abstrait de bien ou de valeur, le simple fait de choisir implique une évaluation, puisque ce qui nous fait choisir une option plutôt qu'une autre est une certaine compréhension que cette option est meilleure. Conceptuellement, le processus de discrimination, qui est la définition même d'un choix, doit impliquer une évaluation, puisque l'évaluation est le moteur même de ce processus. Sans évaluation, il n'y a pas de discrimination possible. Un choix ne peut pas se faire sur la base de rien. Tout choix implique une évaluation.

Lorsqu'on parle d'évaluation, on revient toujours, en dernier lieu, à la notion de valeur, donc de bien. Que l'on évalue une décision possible en termes de coût et de bénéfice, une personne en termes de qualités et de défauts, une sensation en termes de plaisir et de douleur ou une action morale en termes de bien et de mal, on revient toujours à cette dichotomie entre le bon et le mauvais. On favorise le bon et on rejette le mauvais. L'acte d'évaluer est toujours étroitement lié à la notion de valeur. Une évaluation, dans le contexte d'un choix menant à une action, sert à déterminer qu'une action possible est meilleure qu'une autre. Quand on assigne, par le processus d'évaluation, des attributs tels que meilleur ou pire à des choix possibles, et que ces attributs servent à déterminer le choix que nous faisons, on agit en vue du bien. Lorsqu'on détermine qu'une option possible est la meilleure, c'est parce qu'on la croit, d'une manière ou d'une autre, plus bénéfique que les autres. Quand on choisit cette option-là justement parce qu'on la considère comme meilleure, on agit volontairement dans le but de faire survenir le bien. L'évaluation est donc intrinsèquement liée au bien. Toute évaluation doit nécessairement faire intervenir le concept de bien, puisque c'est toujours par ce concept que la valeur d'une chose est jugée en dernier lieu.

L'argument logique en faveur d'une version sans faille de l'approche évaluative se résume donc comme suit. Toute action intentionnelle implique un choix, sous la forme d'une décision de la part de l'agent d'agir comme il agit, sans quoi son action ne peut pas être appelée intentionnelle. Tout choix implique une évaluation, puisqu'un choix doit nécessairement se faire sur la base d'une évaluation, qui permet de discriminer entre les différentes options possibles afin de n'en choisir qu'une. Toute évaluation doit finalement faire référence au bien, puisque c'est par ce concept que l'on peut juger qu'une chose est meilleure qu'une autre et que c'est une considération pour le bien qui nous fait ultimement préférer ce qui est meilleur. Logiquement, donc, on arrive à la conclusion que toute action intentionnelle implique une recherche du bien. Il ne semble pas y avoir de place pour des exceptions.

Le second argument cherche à démontrer que nos intuitions sont fortement alignées sur les principes proposés par l'approche évaluative, ce qui ajoute de la crédibilité à la théorie. L'explication fournie par l'approche évaluative à propos de la motivation à agir s'accorde avec ce que nous pensons ordinairement. Il est normal de penser que nous sommes motivés à accomplir une action quand nous voyons un certain bien dans celle-ci. Ceci est démontré plus clairement lorsqu'on cherche à comprendre les motivations d'autrui. Quand on veut comprendre pourquoi quelqu'un agit d'une certaine manière, on cherche naturellement à savoir quel bien cette personne peut percevoir dans son action. Parallèlement, quand nous arrivons à discerner en quoi une action particulière semble bonne aux yeux d'une personne, nous croyons avoir suffisamment bien compris sa motivation à l'accomplir. En effet, c'est seulement lorsque nous sommes en mesure de discerner cet élément évaluatif dans la pensée de l'agent que nous avons l'impression profonde d'avoir une compréhension satisfaisante de sa motivation.

Imaginons, par exemple, qu'un de mes amis soit un grand collectionneur de bandes dessinées qui dépense temps et argent pour trouver des bandes dessinées rares à ajouter à sa collection. Moi-même, qui ne suis pas collectionneur, ne perçois pas la même valeur que lui dans cette activité. Il ne me viendrait pas à l'esprit d'agir de cette manière. Malgré tout, puisque je sais mon ami passionné de bandes dessinées et que je comprends qu'il

voit de la valeur dans le fait d'acheter des bandes dessinées, je peux comprendre sa motivation à agir ainsi. Je comprends qu'il achète des bandes dessinées parce que cela représente pour lui un certain bien. Sachant cela, je crois avoir une compréhension suffisante de sa motivation à agir ainsi. Je n'éprouve pas le besoin de chercher plus loin. C'est le fait que je puisse discerner la valeur qu'a une action particulière pour un agent qui rend compréhensible à mes yeux sa motivation à agir ainsi.

Inversement, si nous sommes incapables de discerner la valeur qu'un agent perçoit dans une certaine action, alors sa motivation à l'accomplir nous apparaît mystérieuse et incompréhensible. Nous avons alors le sentiment très clair de ne pas comprendre pourquoi l'agent agit ainsi. Un exemple célèbre développé par John Rawls permet d'illustrer ce point.¹³⁰ Imaginons une personne qui passe tout son temps à compter des brins de gazon, et qui s'adonne à cette activité avec un grand dévouement. La plupart d'entre nous sommes incapables de discerner quelle valeur quelqu'un peut percevoir dans cette activité. Pour cette raison, nous ne comprenons pas pourquoi cette personne agit ainsi. Ses motivations demeurent mystérieuses à nos yeux et nous avons l'impression, en pensant à ce cas, qu'il nous manque une information essentielle pour comprendre pourquoi cet individu choisit de compter les brins de gazon de cette manière. Nous allons même jusqu'à qualifier cette action d'irrationnelle, tout cela parce que le bien que l'agent doit percevoir dans cette activité pour agir ainsi nous échappe.

Le bien est donc l'élément que nous recherchons intuitivement quand nous essayons de comprendre les motivations d'autrui. Pour avoir une compréhension satisfaisante du « pourquoi » d'une action particulière, il est nécessaire de savoir de quelle manière l'action peut sembler bonne à l'agent, quelle valeur il perçoit dans ce qu'il fait. Lorsque nous sommes en mesure de discerner cet élément, nous avons naturellement l'impression d'avoir compris sa motivation. Par contre, quand nous sommes incapables de discerner en quoi une action peut sembler bonne ou valable aux yeux de celui qui l'accomplit, nous sommes alors incapables de comprendre pourquoi il agit ainsi. Sa décision nous apparaît mystérieuse.

¹³⁰ Rawls, 1971, p. 432.

Cet argument démontre que nos intuitions sont fortement alignées avec le principe proposé par l'approche évaluative, soit que nous sommes uniquement motivés à accomplir les actions qui nous semblent bonnes. Que la perception du bien soit l'élément que nous recherchons quand on veut comprendre les motivations d'autrui est un indice que celle-ci joue réellement un rôle dans la motivation à agir et dans le processus de prise de décision. On peut donc affirmer, à la lumière de cet argument, que l'approche évaluative est une théorie intuitive, puisqu'elle est conforme avec la manière dont nous appréhendons normalement la motivation à agir.

Le troisième argument en faveur d'une version sans faille de l'approche évaluative postule que l'explication de notre comportement courant rend nécessaire l'existence de capacités évaluatives opérant à un niveau différent que la délibération ordinaire. Cet argument s'appuie sur un type de cas controversé chez les partisans de l'approche évaluative, les actions irréfléchies, et démontre que même celles-ci surviennent à la suite d'une évaluation, ce qui implique l'existence d'un processus évaluatif capable de gouverner des actions irréfléchies. Cette forme d'évaluation moins consciente, mais plus rapide constitue un point en faveur de l'approche évaluative, puisqu'elle permet d'affirmer que même les actions irréfléchies sont accomplies *sub specie boni*, ce qui serait impossible si la seule capacité d'évaluation à notre portée était la délibération rationnelle.

Si les partisans de l'approche évaluative s'entendent pour dire que les actions pleinement délibérées, accomplies par un agent se trouvant dans un état émotionnel stable et possédant des capacités mentales adéquates sont celles qui ont le plus de chances d'être accomplies en vue du bien, d'autres types d'actions moins exemplaires font l'objet d'un débat. C'est le cas, notamment, des actions irréfléchies, souvent décrites comme des actions faites sans raison. Nous pourrions penser que les actions triviales faites par habitude ou par réflexe ne sont pas faites en vue du bien, parce qu'elles n'ont pas assez d'importance à nos yeux pour mériter un effort de délibération. Par exemple, l'acte de boire spontanément une gorgée d'eau d'un verre posé sur la table ne serait pas perçu

comme résultant d'une quelconque évaluation. C'est quelque chose que l'on fait, tout simplement, sans y penser. Cette position est adoptée par Joseph Raz, partisan d'une version modérée de l'approche évaluative. Selon lui, les actions triviales et irréfléchies, comme dans l'exemple du verre d'eau, ne sont pas explicables par l'approche évaluative.¹³¹ Ces actions ne font pas l'objet d'une évaluation au même titre que les actions pleinement délibérées. Raz est d'avis que la plupart de nos actions quotidiennes, celles que nous accomplissons sans réfléchir, ne sont pas accomplies *sub specie boni*, parce qu'elles ne sont pas accomplies avec une intention indépendante, et sont donc véritablement faites sans raison.

Le problème, si on adopte ce point de vue, est que l'immense majorité des actions que nous accomplissons sans réfléchir sont tout de même des actions cohérentes, bien coordonnées et ayant un sens à nos yeux. Ce sont des actions que nous pourrions facilement expliquer et justifier si nous nous en donnions la peine. Ceci ne cadre pas avec l'idée que ces actions sont accomplies sans évaluation. De telles actions devraient être complètement aléatoires, tantôt utiles, tantôt nuisibles, parfois totalement absurdes et incohérentes, mais les actions irréfléchies ne sont pas ainsi dans les faits. Lorsque je prends spontanément une gorgée d'eau, j'accomplis une action ayant un sens à mes yeux, une action que j'aurais facilement pu évaluer et approuver consciemment. De plus, le geste qui consiste à boire une gorgée d'eau est un mouvement précis, bien coordonné et accompli de façon délibérée. Ce n'est pas par hasard que ma main a agrippé le verre avec suffisamment de force pour le soulever, l'a amené vers moi de manière à ne pas renverser l'eau, et l'a porté à ma bouche en l'inclinant juste assez pour me permettre d'en boire. Il est indéniable que l'acte de boire de l'eau n'est pas un geste aléatoire, mais que comme l'immense majorité des actions intentionnelles irréfléchies, il répond à un certain critère de cohérence et d'utilité.

Quiconque voudrait soutenir l'argument que les actions irréfléchies ne sont pas le résultat d'une évaluation et ne tendent pas vers le bien devrait être en mesure d'expliquer comment il serait possible que je prenne une gorgée d'eau sans évaluer d'aucune manière

¹³¹ Raz, 2008, pp. 9-15.

qu'il s'agit d'une bonne action, alors qu'il ne m'arrive jamais de prendre spontanément une gorgée de vinaigre (tout en sachant qu'il s'agit de vinaigre) ou de faire spontanément quoi que ce soit d'autre qui m'apparaît absolument nuisible ou absurde. Si les actions irréfléchies étaient véritablement accomplies sans évaluation, ne devrait-on pas trouver que nous accomplissons, en proportion au moins égale aux actions sensées et utiles, des actions complètement absurdes et incohérentes? Or, cela n'est pas le cas. L'immense majorité des actions que nous accomplissons sans réfléchir se trouvent à être des actions sensées que nous pourrions facilement expliquer et justifier. Il semble donc que nous soyons tout à fait capables d'accomplir des actions bonnes et utiles sans réfléchir ouvertement à ce que l'on fait.

Un exemple particulièrement intéressant pour illustrer ce point est le cas des actions intentionnelles faites par réflexe. Une action faite par réflexe est une action rapide et instinctive survenant à la suite d'un changement soudain de situation. Nous sommes souvent surpris par la rapidité et la précision de nos propres réflexes. Imaginons que j'accroche accidentellement le verre d'eau posé sur la table et que mon geste le précipite vers le sol. En voyant ce qui est en train de se produire, j'ai le réflexe de rattraper le verre d'eau avant qu'il ne tombe, et de le remettre sur la table. Sans réfléchir, j'ai positionné ma main de manière à empêcher le verre de tomber. Cette action faite par réflexe, bien qu'elle ait été accomplie sans réfléchir, doit être considérée comme ayant été faite *sub specie boni*. Pourquoi? Simplement parce que l'acte de rattraper le verre d'eau est une action ayant de la valeur à mes yeux. En fait, parmi toutes les actions possibles au moment où le verre est en train de tomber, celle de le rattraper est sans aucun doute la meilleure possible, et c'est justement celle-là que j'accomplis sans réfléchir. Ce n'est pas une coïncidence qu'un réflexe me fait agir ainsi. Peu importe ce qui a pu déclencher le réflexe, il est nécessaire qu'une certaine évaluation ait eu lieu pour me permettre de reconnaître qu'il valait mieux rattraper le verre d'eau que de le laisser tomber. Nous devons donc admettre que cette action, faite par réflexe, découle d'un certain jugement de valeur.

Si nos actions irréflechies ne sont pas des actions absurdes et incohérentes, mais qu'au contraire elles ont un sens et une utilité, et que l'exemple des actions faites par réflexe permet de démontrer qu'au moins certaines actions irréflechies doivent obligatoirement être motivées par une évaluation, alors nous devons admettre que les actions irréflechies peuvent, au même titre que les actions réfléchies, être gouvernées par un principe évaluatif. Or, puisque les actions irréflechies sont, par définition, des actions au sujet desquelles on ne délibère pas, nous devons admettre l'existence de certains processus évaluatifs opérant à un niveau mental différent que la délibération consciente. De tels processus évaluatifs doivent permettre de guider notre conduite dans les cas où nous ne nous servons pas de notre faculté de délibération rationnelle, qui est trop lente pour nous permettre d'agir rapidement quand la situation l'exige. Il faut donc admettre l'existence de cet élément nécessaire pour que l'approche évaluative puisse véritablement prétendre être une théorie sans faille : une capacité évaluative primaire, différente de celle qu'on appelle la raison. Avec cet élément, l'approche évaluative a tous les outils nécessaires pour expliquer aussi bien les actions pleinement délibérées que la multitude d'actions irréflechies que nous accomplissons tous les jours. Elle peut véritablement prétendre que toutes les actions intentionnelles, sans exception, sont accomplies *sub specie boni*.

Nous ne prendrons pas position dans ce travail sur la question de savoir si cette capacité d'évaluation est cognitive ou non cognitive. Il n'est pas clair que cette question est réellement pertinente dans le cas qui nous concerne. L'important est de reconnaître l'existence de cette capacité d'évaluation primaire et de l'incorporer dans une version sans faille de l'approche évaluative. La théorie doit donc tenir compte d'au moins deux formes distinctes d'évaluation : les jugements de valeur délibérés et l'évaluation primaire qui s'opère à un niveau mental plus instinctif. Il n'est pas impossible, toutefois, que la différence entre ces deux capacités en soit une de degré plutôt que de nature, et qu'elles soient toutes les deux le résultat d'une seule et même propension à rechercher le bien. Cette question mérite une recherche approfondie qui dépasse le cadre de ce travail. Quoi qu'il en soit, une forme d'évaluation telle que celle décrite ci-dessus est absolument nécessaire pour que l'approche évaluative puisse prétendre être sans faille. Sans l'existence d'une capacité évaluative primaire qui ne nécessite pas de capacités

intellectuelles développées pour pouvoir discriminer entre le bon et le mauvais, une théorie sans faille n'a pas les moyens de se défendre contre les multiples objections présentées par ses détracteurs.

3.2 Contre les valeurs anomiques

En adoptant une version sans faille de l'approche évaluative, nous affirmons que chaque action intentionnelle sans exception est motivée par une évaluation, donc par un jugement de valeur. Ceci laisse toutefois la porte ouverte à l'idée que d'autres « valeurs » que le bien puissent motiver. Le mal, par exemple, pourrait-il être une valeur motivant l'action au même titre que le bien? Est-il possible, comme le pense Joseph Raz, d'être motivé par des valeurs anomiques, c'est-à-dire d'être motivé à agir en vue du mal? Dans cette section, nous défendrons l'argument selon lequel il est impossible d'être véritablement motivé par le mal, et que la seule valeur capable de motiver est le bien.

Pour bien comprendre l'argument contre l'idée que le mal peut motiver, il est nécessaire de dissocier complètement notre concept de « mal » de ce qui constitue le mal moral. Dans notre conception ordinaire, le mal est encore plus étroitement associé avec la moralité que le bien. C'est pourquoi il est important de souligner que le mal auquel nous faisons référence ici n'a rien à voir avec la moralité. Il réfère à tout ce qui peut apparaître mauvais ou indésirable aux yeux d'un agent. Il serait faux d'affirmer qu'il est impossible d'être motivé à accomplir une action moralement mauvaise. Nous avons brièvement mentionné précédemment l'exemple d'un voleur souhaitant dérober un objet précieux. Il sait que son action est moralement mauvaise, mais il désire suffisamment l'objet en question pour que l'acte de le voler lui apparaisse bon. Il est possible de percevoir des actions particulières comme étant bonnes même lorsqu'il s'agit d'actions moralement mauvaises ou encore d'actions autodestructrices. Il est également possible d'avoir une compréhension oblique d'une action comme étant mauvaise, tout en ayant une compréhension directe de son bien, comme nous l'avons vu à la section 2.1. Ce qui semble impossible, cependant, c'est d'avoir une compréhension d'une chose comme étant à la fois bonne et mauvaise, désirable et indésirable, toutes choses considérées.

L'approche évaluative postule qu'être motivé à agir signifie être attiré, en quelque sorte, par l'idée d'accomplir une action particulière. Il ne fait aucun doute que nous soyons généralement attirés par ce qui nous semble désirable ou attrayant. Quand une chose nous apparaît bonne, nous désirons cette chose et nous sommes attirés par elle. Cette forme d'attraction est caractéristique de l'attitude de désirer. Au contraire, lorsqu'une chose nous apparaît indésirable, elle apparaît répugnante et génère en nous une attitude de répulsion ou d'évitement. Cette attitude est contraire à celle qui est caractéristique au désir, et ces deux attitudes sont conceptuellement liées aux notions de désirable et d'indésirable. Tout comme il serait absurde de dire que nous sommes repoussés par ce que nous désirons, il est absurde d'affirmer que nous pouvons être attirés par ce qui nous semble indésirable. C'est cette attitude naturelle vis-à-vis de l'indésirable qui nous permet d'affirmer que l'on ne peut pas être motivé par des valeurs anomiques, c'est-à-dire par le mal *simpliciter*. En effet, pour imaginer que la perception du mal puisse motiver à agir, il faudrait imaginer que la caractéristique qu'a un objet d'être repoussant puisse elle-même générer en nous une attraction pour cet objet, ce qui est un point de vue difficile à défendre. Pour cette raison, nous sommes en désaccord avec une position comme celle de Raz, qui affirme que la perception du mal peut parfois motiver au même titre que la perception du bien.

Imaginons, par exemple, que je sois assis devant un foyer et que, dans un moment contemplatif, je considère la possibilité de mettre ma main dans le feu. J'imagine en détail les conséquences probables de cette action. J'imagine la douleur intense de la brûlure et la possibilité de subir des blessures permanentes. Je considère également le caractère tout à fait absurde et stupide d'un tel geste. En considérant ces différentes conséquences relatives à la possibilité de mettre ma main au feu, je suis profondément répugné par celles-ci. L'idée de m'infliger à moi-même des blessures sérieuses de la manière la plus stupide qui soit m'apparaît si mauvaise qu'elle est accompagnée d'un intense sentiment de répulsion, d'un désir d'éviter ces conséquences à tout prix. C'est sur ce sentiment de répulsion, plutôt que sur l'exemple lui-même, qu'il faut porter notre attention ici. Ce sentiment accompagne naturellement la contemplation de choix qui nous semblent mauvais. Pour croire que l'apparence de mal peut motiver à agir, il faudrait

croire que, justement, ce sentiment de répulsion que nous avons envers ce qui nous semble indésirable serait en mesure de nous motiver à accomplir cette action indésirable. Ceci est difficile à imaginer, mais c'est justement la vérité qu'il faudrait admettre si nous voulions avancer que le mal *simpliciter* peut motiver dans certaines circonstances.

Ainsi, concevoir que le mal puisse motiver requiert de nous d'admettre une idée fortement contre-intuitive, soit que le sentiment de répugnance qui accompagne typiquement la contemplation d'actions qui nous semblent mauvaises pourrait avoir pour effet de nous motiver à poser ces actions. Ceci ne revient pas à dire que ce sentiment ne peut pas être vaincu par une considération plus forte pour le bien. Il est possible d'avoir le désir d'accomplir une action qui nous semble mauvaise, mais ce désir n'est pas motivé par la répugnance elle-même, mais plutôt par des considérations différentes qui confèrent à l'action une certaine valeur à nos yeux en dépit de cette répugnance. Toutefois, si on pense strictement au mal *simpliciter*, non teinté par des considérations diverses pour le bien, et qu'on se demande si la perception de ce mal peut motiver à agir, il m'apparaît clairement que la réponse est non. La version de l'approche évaluative défendue dans cet ouvrage devra donc rejeter la possibilité que des valeurs anomiques puissent motiver l'agent à accomplir ce qui lui semble mal.

3.3 Contre une théorie non séparatiste

Une autre source de divergence possible chez les partisans de l'approche évaluative se situe au niveau de la portée de la théorie. La version radicale de l'approche évaluative proposée par Sergio Tenenbaum est une théorie non séparatiste qui postule une double implication entre motivation et évaluation, affirmant que tous les états motivationnels sont des états évaluatifs et que tous les états évaluatifs sont des états motivationnels.¹³² Selon cette version, non seulement la motivation implique l'évaluation au sens où chaque action intentionnelle est motivée par une évaluation, mais l'évaluation implique également la motivation, puisque toutes les fois où nous évaluons qu'une chose est bonne, nous sommes motivés à agir en ce sens. Joseph Raz, pour sa part, défend une

¹³² Tenenbaum, 2007, pp. 2-3.

version séparatiste qui affirme que la motivation n'implique pas l'évaluation, c'est-à-dire qu'on peut être motivé à accomplir des actions qui ne nous semblent pas bonnes. Toutefois, en prenant position en faveur d'une version sans faille, nous avons déjà jugé que la motivation implique l'évaluation. Il reste à savoir si, comme le pense Tenenbaum, l'évaluation implique également la motivation, ou si au contraire il est possible d'évaluer sans être motivé à agir. Dans cette section, nous argumenterons contre l'idée que l'évaluation implique la motivation et nous nous positionnerons en faveur d'une théorie séparatiste qui admet l'existence d'états évaluatifs non motivationnels.

Les arguments présentés jusqu'ici dans ce travail étayent l'idée que la motivation implique l'évaluation, que chaque action intentionnelle est accomplie en vue du bien. Cette hypothèse a été étudiée et analysée sous tous ses angles, les objections contre elle ont été discutées minutieusement, et nous avons maintenant de bonnes raisons de penser que cette idée est vraie. Toutefois, l'hypothèse inverse, à savoir que toutes les fois où l'agent évalue qu'une chose est bonne, il est motivé à agir, n'a pas été défendue ici, et nous devons nous garder de la considérer nécessairement vraie. À première vue, cette position semble difficile à défendre. On peut penser à de multiples scénarios plausibles où nous jugeons qu'une chose est bonne, sans que cette évaluation s'accompagne d'une motivation à agir. Il nous arrive souvent de porter des jugements de valeur sur des choses qui sont totalement hors de notre contrôle et qui n'appellent pas à une action de notre part. Je peux, par exemple, penser que c'est une bonne chose qu'il existe une station spatiale internationale. Cette pensée n'appelle pas à une action de ma part, elle ne me motive pas à faire quoi que ce soit. Cette évaluation n'est donc pas liée à une motivation à agir. L'évaluation que je porte envers la station spatiale internationale est passive, il s'agit d'un jugement contemplatif.

Il serait possible de rétorquer que des jugements de valeur comme celui-ci motivent bel et bien à agir, puisqu'ils me prédisposent à supporter l'existence de la station spatiale et à encourager des projets futurs du même genre, et que cette prédisposition constitue une forme très primaire d'action, une sorte d'approbation mentale continue. Toutefois, il est douteux que tous les jugements contemplatifs doivent nécessairement être associés avec

ce genre de motivation. Il serait exagéré de dire qu'un simple moment contemplatif, où je pense pendant quelques secondes à la station spatiale internationale avant que mon attention soit dirigée sur autre chose, doit nécessairement faire naître en moi une motivation à agir. Au contraire, cette très simple évaluation ne semble générer aucune motivation en moi. La valeur que je perçois dans l'existence de la station spatiale n'est pas une valeur devant être réalisée par mon action, c'est une valeur qui existe quoi que je fasse et que je ne fais que contempler.

D'autres exemples peuvent illustrer ce point de manière plus évidente. Imaginons que je possède un billet de loterie pour le tirage du lendemain. Je me dis alors que ce serait une bonne chose pour moi de gagner le gros lot. Cette évaluation ne me motive pas à acheter un billet, puisque j'en ai déjà un. Elle ne me pousse pas à vérifier les numéros gagnants, puisqu'ils ne sont pas encore tirés. Elle ne me motive à rien, puisqu'il n'existe aucune action que je puisse accomplir présentement et qui ferait survenir la valeur que je perçois. La situation est entièrement hors de mon contrôle. Le fait de penser qu'il serait une bonne chose que je gagne le gros lot, alors que j'ai déjà un billet entre les mains, est une évaluation contemplative, pas un motivateur à l'action. Encore un autre exemple d'évaluation contemplative est la remémoration d'événements du passé. Quand nous pensons au passé, nous portons souvent des jugements de valeur sur nos propres actions et sur celles des autres. De telles évaluations ne nous engagent pas à agir, cependant, parce que la valeur perçue dans les événements du passé n'est pas présentement réalisable par l'action. Elle a déjà été réalisée. Il est possible que parfois, penser à des événements du passé nous motive à accomplir une action dans le présent, mais ce n'est pas toujours le cas, et cette motivation est souvent causée par des considérations pour le présent et le futur. En fin de compte, les événements du passé sont simplement une autre chose sur laquelle on peut porter un jugement de valeur sans pour autant être motivé à agir.

Nous pouvons donc affirmer, sans prendre trop de risques, que certaines évaluations ou apparences de bien ne motivent pas à l'action. Les états évaluatifs ne s'accompagnent pas toujours, comme le pense Tenenbaum, d'états motivationnels conséquents. Mais alors, si

certaines perceptions de valeur motivent et d'autres pas, lesquelles motivent et pourquoi? Tenenbaum affirme, avec raison, que le bien est le but constitutif de l'action.¹³³ Quand on agit, c'est en vue de faire survenir un certain bien. À partir de là, il est logique d'affirmer que les perceptions de valeur capables de motiver sont celles qui présentent le bien comme devant être réalisé par l'action. Selon ce modèle, la motivation à agir serait donc la motivation à faire survenir un bien réalisable par l'action. Les exemples que nous avons examinés dans cette section sont des cas où le bien est certes perçu par l'agent, mais où il n'est pas considéré par lui comme devant être réalisé par son action. Il est naturel, alors, que dans ces cas-là la perception du bien ne motive pas à agir, puisqu'elle ne saurait guider l'agent vers une action particulière qui soit en mesure de faire survenir le bien en question. Une valeur perçue comme n'étant pas réalisable par l'action ne motive pas à agir.

Dans cette section, nous avons défendu une version séparatiste de l'approche évaluative. Selon cette version, la motivation implique l'évaluation, mais pas l'inverse. Il existe des états évaluatifs qui ne sont pas motivationnels. La différence entre une version séparatiste comme celle-ci et une version non séparatiste de l'approche évaluative est que la version séparatiste est à la fois moins radicale et plus plausible que la version non séparatiste. Elle est moins radicale parce qu'elle affirme moins de choses. Elle affirme qu'il existe une simple implication entre motivation et évaluation, mais pas une double implication. En termes logiques, ceci constitue une différence majeure entre les deux versions de la théorie. En plus d'être moins radicale, la version séparatiste est plus plausible que la version non séparatiste, puisque l'hypothèse postulée par cette dernière, à savoir que l'évaluation implique la motivation, est difficile à défendre compte tenu des contre-exemples qui démontrent que nous pouvons porter des jugements de valeur sans pour autant être motivés à accomplir une action particulière en réponse à ces jugements. Pour cette raison, nous concluons qu'une version optimale de l'approche évaluative doit être une version séparatiste qui affirme uniquement que la motivation implique l'évaluation.

¹³³ *Ibid.*, p. 6.

CONCLUSION

Dans le cadre de cette recherche, nous avons présenté une théorie de l'action intentionnelle, l'approche évaluative, défendu celle-ci contre plusieurs objections, puis dégagé, au terme de notre recherche, la version de cette théorie qui nous semble la meilleure. Voici donc, sommairement, la théorie que nous endossons. Nous adhérons, premièrement, au principe général de l'approche évaluative, qui est que les actions intentionnelles sont accomplies en vue du bien (*sub specie boni*). En disant cela, nous maintenons qu'une certaine forme d'évaluation motive l'agent à accomplir chacune de ses actions et que cela constitue une poursuite du bien. Pour être explicable par l'approche évaluative, une action doit être intentionnelle (l'agent doit l'accomplir volontairement) et intelligible (une réponse satisfaisante au « pourquoi? » d'Anscombe doit pouvoir être donnée). Dès lors qu'une action satisfait à ces critères, elle est accomplie en vue du bien. Il n'y a pas d'exceptions possibles à ce principe. Des valeurs anomiques, telles que le mal, ne peuvent pas motiver à agir. Seul ce qui est perçu comme étant bon, ou valable, peut être désiré. La théorie que nous adoptons est toutefois une théorie séparatiste. Elle affirme que bien que chaque action intentionnelle soit motivée par une évaluation, l'inverse n'est pas vrai et il est possible d'évaluer qu'une chose est bonne sans être motivé à agir. Nous avançons l'hypothèse que les seules évaluations capables de motiver sont celles dont le bien est compris comme étant réalisable par l'action, puisque toute action a pour but de faire survenir le bien.

Cette théorie est plus radicale que la version proposée par Joseph Raz. En termes syllogistiques, celui-ci affirme que certaines actions intentionnelles ne sont pas accomplies en vue du bien. Nous affirmons que toutes les actions intentionnelles sont accomplies en vue du bien. La théorie que nous proposons est cependant moins radicale que celle présentée par Sergio Tenenbaum. Bien que les deux théories s'accordent pour dire que toutes les actions intentionnelles sont accomplies en vue du bien, Tenenbaum ajoute que toutes les perceptions de valeur motivent à agir, tandis que nous affirmons que certaines perceptions de valeur ne motivent pas à agir. Le fait que notre version de l'approche évaluative demeure « sans faille » est sans doute suffisant pour dire qu'elle est

une théorie radicale, mais en réalité elle couvre davantage un juste milieu entre les deux extrêmes possibles de l'approche évaluative (ou « Guise of the Good Theory »).

Notre version de l'approche évaluative résiste bien aux diverses objections adressées à la théorie en général. Premièrement, les contre-exemples présentés en tant qu'objections à la théorie s'avèrent peu convaincants. Dans la plupart des cas, il s'agit de quelqu'un dans un état d'esprit anormal accomplissant une action que la plupart des gens désapprouveraient. Les détracteurs de l'approche évaluative utilisent ces contre-exemples pour affirmer que dans ces cas-là, l'agent ne pouvait pas voir du bien dans son action, parce que nous n'en voyons pas nous-mêmes. Il est plus raisonnable, à mon avis, de penser que lorsque nos états mentaux changent, notre perception du bien peut également changer, mais pas notre propension à vouloir agir en vue du bien. Une telle explication peut nous permettre de comprendre la cause de motivations déviantes, puisque l'on peut dire, par exemple, que des états mentaux tels que la colère et la dépression font soudainement apparaître certaines actions destructrices comme étant bonnes, alors que normalement nous les considérons mauvaises. Ceci est plus plausible que de dire que parfois, l'agent n'agit pas *sub specie boni*, mais qu'il agit soudainement selon un principe différent, ou encore sans raison.

Agir *sub specie boni* ne nécessite pas de maîtriser le concept de bien, ni de posséder des capacités intellectuelles sophistiquées. Il est tout à fait possible, selon notre point de vue, que même les animaux agissent en prenant le bien comme but. La propension naturelle que nous avons à rechercher le bien ne complique pas notre prise de décision. Au contraire, elle la simplifie. Dans bien des cas, la perception de valeur se présente à nous et génère une motivation qui est naturellement forte. Il n'est alors pas tant question de décider ce que l'on veut faire que de découvrir ce que l'on veut faire. Si les animaux ne font qu'agir en suivant les perceptions de valeur qui se présentent à eux, ils peuvent agir efficacement sans avoir à analyser la situation en profondeur. Si, au contraire, ils n'agissent pas en prenant la perception du bien comme guide de leur action, alors il faut leur attribuer des capacités analytiques plus poussées qui puissent compenser l'absence de cette motivation simple à rechercher le bien. Ainsi, l'approche évaluative ne requiert

pas qu'un agent possède des capacités intellectuelles développées pour agir en vue du bien. Au contraire, le désir du bien est un outil permettant de simplifier la prise de décision, et donc de compenser efficacement le manque de capacités analytiques plus avancées, notamment chez les animaux.

Nous avons vu également que la contrainte de la croyance, imposée par Setiya, n'est pas une menace sérieuse pour la théorie. L'accusation adressée à l'endroit de l'approche évaluative, selon laquelle elle ne permettrait pas de connaître nos propres actions, repose sur une interprétation sémantique douteuse. Que nous soyons ou non d'accord avec l'idée que nous avons une connaissance sans observation de nos actions, l'approche évaluative n'est pas vraiment mise en danger par la contrainte de la croyance, puisqu'il n'est pas réaliste de penser qu'une théorie qui affirme que nous agissons dans le but de faire survenir le bien serait par ailleurs incapable d'affirmer que nous connaissons nos propres actions.

Nous avons finalement analysé les problèmes inhérents à une théorie qui tenterait d'expliquer les raisons d'agir sans faire référence à une quelconque évaluation de la part de l'agent. Une telle théorie serait extrêmement limitée dans son explication du pourquoi de l'action, et n'offrirait pas une compréhension satisfaisante des raisons d'agir. L'évaluation est un élément nécessaire à la compréhension de nos raisons d'agir. Sans cet élément, les explications du pourquoi de l'action sont vides de sens. Le bien et la valeur doivent faire partie de toute explication pertinente de la motivation à agir et on ne peut pas ignorer ces éléments sans faire face à des problèmes insurmontables.

L'adoption de l'approche évaluative aurait des implications importantes pour la philosophie de l'action. C'est une théorie fonctionnelle qui peut contribuer à faire progresser la connaissance de l'action intentionnelle en nous permettant de déduire de nouvelles informations à partir de données existantes. Si une personne accomplit intentionnellement une action, nous pouvons en déduire qu'elle percevait de la valeur dans cette action. Nous pouvons alors chercher à savoir quelle est cette valeur, tout en étant sûrs qu'il y a bel et bien quelque chose à trouver et que l'action n'est pas survenue

spontanément, sans raison. Inversement, si nous connaissons quelles choses apparaissent bonnes à un agent, nous pourrions alors mieux estimer comment il va agir, et nous saurons pourquoi il agit ainsi. L'approche évaluative permet d'élucider certains éléments qui demeureraient mystérieux dans notre compréhension de l'intention humaine. Elle permet d'avoir une compréhension satisfaisante du « pourquoi » des actions intentionnelles.

L'approche évaluative ouvre aussi la porte à la possibilité d'intégrer la perception de valeur et les intentions dans une chaîne causale, ce qui rendrait possible une compréhension psychologique du déterminisme. En nous permettant de voir comment nos propres actions sont causées par une perception de valeur, l'approche évaluative peut nous aider à comprendre comment nous pouvons nous-mêmes faire partie, en tant qu'êtres pensants, d'une chaîne causale universelle, et réconcilier l'existence de notre conscience et de notre volonté avec une vision déterministe de l'univers. L'approche évaluative a donc beaucoup à offrir, si on l'accepte pour vraie, non seulement à la philosophie de l'action, mais aussi à notre compréhension plus générale du monde et de notre place dans celui-ci.

BIBLIOGRAPHIE

- Anscombe, G. E. M., 1963. *Intention*, 2nd ed. Ithaca, N.Y.: Cornell University Press.
- Aristote, *Éthique à Nicomaque*, Bodéüs, R. (trad.), 2004. Paris, GF Flammarion.
- Bermudez, J.L., 1998. *The Paradox of Self-Consciousness*, MIT Press.
- Davidson, Donald, 1980. *Essays on Actions and Events*, Oxford: Clarendon Press.
- Friedrich, Daniel, 2012. « The alluringness of desire », *Philosophical Explorations: An International Journal for the Philosophy of Mind and Action*, pp. 291-302.
- Hawkins, Jennifer, 2008. « Desiring the Bad under the Guise of the Good », *The Philosophical Quarterly*, Vol. 58, No. 231, pp. 244-264.
- Boyle, Matthew et Douglas Lavin, 2010. « Goodness and Desire », dans Sergio Tenenbaum (ed.), *Desire, Practical Reason, and the Good*. New York: Oxford University Press, pp. 158–99.
- Moore, G. E., 1903. *Principia Ethica*, Cambridge: Cambridge University Press.
- Moss, Jessica, 2010. « Aristotle's Non-Trivial, Non-Insane View that Everyone Always Desires Things under the Guise of the Good », dans Sergio Tenenbaum (ed.), *Desire, Practical Reason, and the Good*. New York: Oxford University Press, pp. 65-81.
- Platon, *Oeuvres complètes*, Brisson, L. (dir.), 2008. Paris, Garnier-Flammarion.
- Rawls, John, 1971. *A Theory of Justice*, Belknap Press.
- Raz, Joseph, 1999. *Engaging Reason: On the Theory of Value and Action*, Oxford: Oxford University Press.
- Raz, Joseph, 2008. « On the Guise of the Good », *Oxford Legal Studies Research Paper* No. 43/2008.
- Raz, Joseph, 2015. « The Guise of the Bad », *Columbia Public Law Research Paper* No. 14-457.
- Setiya, Kieran, 2007. *Reasons without Rationalism*, Princeton: Princeton University Press.
- Setiya, Kieran, 2010. « Sympathy for the Devil », dans Sergio Tenenbaum (ed.), *Desire, Practical Reason, and the Good*. New York: Oxford University Press, pp. 82–111.
- Stocker, Michael, 1979. « Desiring the Bad: An Essay in Moral Psychology », *The Journal of Philosophy*, vol. 76, no. 12, pp. 738-753.
- Sumner, L.W., 1996. *Welfare, happiness, and ethics*, Oxford: Clarendon Press.
- Tenenbaum, Sergio, 2007. *Appearances of the Good: An Essay on the Nature of Practical Reason*, New York: Cambridge University Press.

- Tenenbaum, Sergio, 2012. « Knowing the Good and Knowing What One Is Doing », *Canadian Journal of Philosophy*, suppl. vol. 35 (Belief and Agency), pp. 91–117.
- Tenenbaum, Sergio, 2013. « Guise of the Good », *The International Encyclopedia of Ethics*, Blackwell Publishing.
- Velleman, David, 1992. « The Guise of the Good », *Noûs*, vol. 26, no. 1, pp. 3–26.
- Voigt, Katja Maria, 2010. « Desiring the Good: A Socratic Reading of Aristotle », Non-publié.