

La taxinomie des émotions: désunité contre unité*

Guillaume Bard**

Résumé

Dans cet essai, j'expose d'abord la thèse désunitaire de Paul Griffiths, pour qui les émotions se divisent en deux genres naturels fort distincts, et ce, bien qu'on les regroupe communément sous le même genre nominal. Ainsi, pour Griffiths, les émotions sont soit des émotions de base, relevant du réflexe ou de l'instinct, soit des émotions cognitives, requérant une forme de délibération, de jugement, de faculté morale. J'évoque ensuite les différentes avenues que peuvent emprunter les défenseurs d'une thèse unitaire qui, à l'encontre de celle de Griffiths, proposerait de théoriser sous un modèle unique l'ensemble des émotions. Parmi celles-ci, j'examine la théorie des émotions exposée par Jenefer Robinson, en soulignant à la fois son rejet des positions cognitivistes et son désir de conserver certains points forts de ces mêmes positions. Après avoir mis en relief le caractère affectiviste de sa position, je relèverai chez Robinson un problème méthodologique susceptible d'affaiblir passablement l'argument selon lequel il s'agit d'une théorie susceptible de rendre compte de l'ensemble des émotions.

1. Introduction : un horizon dualiste

Le contraste est vif entre, d'une part, les émotions qui semblent relever du réflexe ou de l'instinct et, d'autre part, les émotions qui

* La version initiale de cet essai a été rédigée à l'hiver 2009 dans le cadre d'un séminaire intitulé Problèmes de métaéthique, sous la direction de Christine Tappolet.

** L'auteur est étudiant à la maîtrise en philosophie (Université de Montréal).

requièrent une forme de délibération, de jugement, de faculté morale¹. D'un côté, on constate l'existence de réactions émotives dont on croit, eu égard à leur spontanéité, qu'elles ne peuvent pas être le résultat d'une longue et mûre réflexion. Ainsi, dans l'exemple typique évoqué par William James, une personne qui se retrouve nez à nez avec un ours réagira promptement, fuira le danger, sera prise de peur, et ce, sans avoir procédé à une évaluation cognitive de la situation. D'un autre côté, certaines émotions apparaissent inextricablement liées à la réflexion, au sens où ces émotions sont causées par des évaluations cognitives. Une émotion comme la honte suppose, chez celui qui la ressent, la présence d'un certain type de jugement, par exemple la croyance que sa situation est déshonorante. On comprendrait mal que quelqu'un éprouve de la honte s'il juge, en même temps, que son état ou ses actions n'ont rien de déshonorant, ni à ses yeux, ni aux yeux de son entourage.

Les lignes de fracture entre les deux types d'émotions semblent nombreuses : nature/culture, corps/esprit, passion/raison, affect/cognition, etc. Ce dualisme est-il inévitable? Doit-on penser que les émotions instinctives et irréfléchies, comme la peur de l'ours, ne pourraient pas légitimement être considérées comme faisant partie de la même catégorie que les émotions plus cognitives, requérant une part de compréhension ou de conscience face à la situation qui les provoque? Peut-on développer une théorie commune qui puisse aussi bien rendre compte des émotions de base que des émotions cognitives? Une émotion complexe peut-elle être réduite à une émotion de base, ou doit-elle plutôt demeurer farouchement indépendante de toute tentative d'assimilation?

Dans cet essai, j'exposerai d'abord la thèse de Paul Griffiths, pour qui les émotions se divisent en deux *genres naturels* distincts, bien qu'on les regroupe, communément, sous le même genre nominal. J'évoquerai ensuite les différentes avenues que pourraient emprunter les défenseurs d'une thèse unitaire, c'est-à-dire une thèse qui, à

¹ Tout au long de cet essai, j'utiliserai principalement les expressions « émotions de base » et « émotions cognitives supérieures » pour référer aux deux types d'émotions. On considérera toutefois les termes « programmes affectifs » et « émotions primaires » comme quasi-synonymes de l'expression « émotion de base ». Les « émotions secondaires » renvoient quant à elles aux émotions cognitives supérieures.

l'encontre de celle de Griffiths, proposerait de théoriser sous un modèle unique l'ensemble des émotions. J'évaluerai par la suite la théorie des émotions exposée par Jenefer Robinson, en soulignant à la fois son rejet des positions cognitivistes et son désir de conserver certains points forts de ces mêmes positions. Je mettrai en relief le caractère unitaire et affectiviste de sa théorie, ainsi que son positionnement vis-à-vis des objections formulées par Griffiths. Enfin, je relèverai, chez Robinson, une faiblesse de nature essentiellement méthodologique, susceptible d'assombrir passablement sa position.

2. La théorie des genres naturels et la taxinomie des émotions

Paul Griffiths soutient que sous le vocable « émotion » sont subsumées au moins deux classes différentes d'objets², l'une référant aux émotions de base, et l'autre, aux émotions cognitives supérieures. La question qui sous-tend la thèse de Griffiths est la suivante : le concept vernaculaire d'émotion renvoie-t-il à un genre unifié, et si oui, s'agit-il d'un genre naturel? Ou doit-on plutôt penser que les deux classes d'émotions sont bel et bien distinctes, et que leur unification sous un même genre est nécessairement artificielle? Face à cette

² Griffiths croit par ailleurs qu'un troisième phénomène serait communément assimilé — à tort — au terme émotion. Il s'agirait en fait d'un état psychologique *semblable* à l'émotion, notamment en ce qui a trait à son expression comportementale et à son caractère apparemment involontaire. Ce phénomène, que Griffiths nomme « *disclaimed action emotion* », est caractérisé par l'adoption inconsciente, chez un agent, d'un rôle socialement construit. Ainsi l'agent, sous la pression de la conformité (pression qui n'est explicite ni pour lui ni pour la société qui l'exerce à son égard), adopte un comportement déterminé, caractérisé par une apparente perte de contrôle vis-à-vis ses actions. On supposera donc l'existence d'une émotion interférant dans les motivations de l'agent, afin de rendre compte de l'irrationalité du comportement. Pour Griffiths, ce troisième phénomène doit être traité comme un cas parasitaire et imitatif de l'émotion véritable. L'émotion postulée à des fins explicatives n'est tout simplement pas réelle. (Paul E. GRIFFITHS, *What Emotions Really Are: The Problem of Psychological Categories*, Chicago, University of Chicago Press, 1997, p. 245).

interrogation, Griffiths défend la thèse de la *désunité*, selon laquelle les phénomènes regroupés sous le concept usuel d'émotion forment une classe hétérogène, trop hétérogène pour répondre aux exigences liées à la notion de genre naturel. Cette désunité suggérerait que soient étudiées distinctement, et dans le respect de leurs différences, les émotions de base et les émotions cognitives supérieures.

Mon objectif ici n'est pas d'exposer en détail l'argumentation de Paul Griffiths en ce qui a trait à la validité de la notion de genre naturel. Toutefois, dans la mesure où Griffiths réfute l'idée selon laquelle les émotions constituent une telle catégorie, il importe de résumer quelques-uns des arguments sur lesquels Griffiths fonde ses conclusions.

Il existe plusieurs façons dont on peut classifier les objets (taxinomies), et la division d'une pluralité d'objets en catégories cohérentes dépendra nécessairement de la méthode employée (méthode dont le choix pourra dépendre, à son tour, du but recherché). Plusieurs découpages différents sont possibles pour les objets d'un même ensemble, et chacun des découpages repose sur une méthode distincte. Par exemple, on pourrait classifier les oiseaux de proie selon qu'ils soient migrateurs ou résidents, mais on pourrait également privilégier une autre méthode, grâce à laquelle ces oiseaux seraient catégorisés en rapaces diurnes ou nocturnes. Évidemment, ce qui ressort d'un tel constat, c'est le caractère hautement arbitraire du choix de la méthode de division. En privilégiant arbitrairement une méthode de division plutôt qu'une autre, on s'expose au risque que les catégories ainsi créées correspondent à des genres purement artificiels.

Toutefois, la notion de genre naturel devrait pouvoir échapper à cette impasse, du moins si on adopte une position réaliste, c'est-à-dire une position selon laquelle la nature impose ses propres catégories : “ Natural kinds [are] categories that supposedly correspond to some real distinctions in nature and around which theories are structured.”³ Dans cette optique, le rôle des concepts serait de circonscrire des catégories réelles. Les divisions naturelles précéderaient — et guideraient — les divisions théoriques : “ The adherents of a scientific theory are committed to some account of the

³ *Ibid.*, p. 171.

ultimate nature of reality. They identify the natural kinds in their theory with constituents of this reality.”⁴

Ce réalisme doit cependant être contrasté avec la position rivale, qui affirme que toute classification de la nature dépend, en quelque sorte, d’un cadre théorique préétabli⁵. En d’autres termes, la nature ne serait pas au fondement de la façon dont nous la catégorisons. La subdivision de la nature en classes serait plutôt le résultat d’une projection de nos concepts sur la nature. Ainsi, la légitimation d’une théorie sur la base du caractère supposément objectif et réel des catégories naturelles risque d’aboutir à une certaine circularité : l’analyse théorique du monde s’appuie sur des catégories qu’elle juge inhérentes au monde, sans s’apercevoir que ces catégories sont en fait projetées *par elle* sur le monde.

The world cannot be clearly separated from the mind that understands it. Categories are dependent in some way on the theoretical concepts that represent them [...] If categories are in part dependent on our theories, then they cannot one-sidedly dictate how our theories should develop⁶.

Ce seraient donc nos concepts qui délimiteraient les catégories naturelles, et non l’inverse. Doit-on en conclure que, n’étant pas héritées de la nature, nos taxinomies seraient nécessairement infondées, ou à tout le moins artificielles? Cette avenue invaliderait, semble-t-il, la notion même de genre naturel.

Face à cette objection, Paul Griffiths croit qu’il est possible de défendre l’idée du genre naturel sans nécessairement souscrire à un réalisme des catégories⁷. Sa démarche consiste à exposer la fonction

⁴ *Ibid.*, p. 174.

⁵ Cette position pourrait s’apparenter à l’affirmation de Goodman, à l’effet que « quoi qu’on ait à décrire, on est limité par les manières de décrire » (Nelson GOODMAN, *Manières de faire des mondes*, Nîmes, Éditions Jacqueline Chambon, 1992, p. 11).

⁶ Paul E. GRIFFITHS, *op. cit.*, p. 174.

⁷ Pour défendre cette thèse, Griffiths s’appuie sur les études de Richard Boyd. (Richard BOYD, « What realism implies and what it does not », *Dialectica*, vol. 43, no 1-2, p. 5-29.)

épistémique des catégories. Certaines, plus que d'autres, auraient un mérite « explicatif », et ce, en vertu de leur projectibilité⁸ :

When organisms are divided into species or substances divided into chemical elements, they are divided in ways that are projectable. The correlations of properties that these categories represent can be relied upon to hold up in new instances. A new organism that passes the test for membership of a species will have the other features characteristic of the species but not specified in the test⁹.

Le caractère unifié et homogène d'une catégorie dépendra de la projectibilité des composantes qui auront été identifiées comme essentielles.

Griffiths demeure sur le terrain anti-réaliste, puisque c'est la *postulation* d'une essence à des fins explicatives qui semble compter, et non sa réalité. Ce qui est postulé comme étant l'essence d'une catégorie n'est, en fin de compte, que le mécanisme causal supposé fournir l'explication la plus probante en ce qui a trait à la cooccurrence des propriétés :

Boyd and Keil use the idea of *causal homeostatis* [...] to capture what makes a category useful. A category brings together a set of objects with correlated properties. The category has causal homeostatis if this set of correlations has some underlying explanation that makes it projectable. A successful category captures what Keil calls a causal homeostatic mechanism—something which means that the correlations can be relied on to hold up in unobserved instances. The search for causal homeostatis explains what has been called *psychological essentialism*. People do not

⁸ Ici, Griffiths fait référence au travail de Frank C. Keil sur la formation et la révision des concepts dans la psychologie développementale. (Frank C. KEIL, *Concepts, Kinds and Cognitive Development*, Cambridge, MIT Press, 1989, 328 p.)

⁹ Paul E. GRIFFITHS, *op. cit.*, p. 174.

simply note the existence of clusters of properties. They postulate a system of underlying causes of the clustering¹⁰.

On peut donc concevoir qu'il existe plusieurs classifications concurrentes, relatives à un même domaine. Une position anti-réaliste des catégories proscrit d'emblée la tentative de privilégier une taxinomie plutôt qu'une autre en invoquant l'existence des divisions inhérentes à la nature. La réfutation du réalisme des catégories n'implique toutefois pas que l'on doive nécessairement souscrire à une forme de relativisme, aux yeux de laquelle toutes les taxinomies se vaudraient, puisque aucune d'elles ne serait fondée sur la réalité. La théorie de l'homéostasie causale fonde la légitimité d'une taxinomie sur le fait qu'un mécanisme causal sous-jacent peut être invoqué pour justifier l'homogénéité des catégories et prédire, dans une certaine mesure, sa généralisation aux cas à venir.

Griffiths illustre son propos à l'aide du contre-exemple suivant : le terme anglais *lilies* réfère à une catégorie non homogène du point de vue homéostatique, qui regroupe sous le même vocable plusieurs liliacées mais aussi certaines fleurs appartenant à d'autres familles.

In the light of the causal homeostasis theory, it is evident that lily is a nominal kind, defined by flower form, rather than a natural kind. [...] The folk category of lilies is of extremely limited use for explanation and induction. It collects no cluster of properties beyond a certain vague visual gestalt¹¹.

Il en irait de même pour les émotions.

C'est en vertu de leur *passivité*¹² que les émotions de base et les émotions cognitives supérieures se voient regroupées au sein d'une

¹⁰ *Ibid.*, p. 188.

¹¹ *Ibid.*, p. 191.

¹² Par *passivité*, on doit entendre une certaine interférence des émotions dans le flux des actions planifiées à long terme (*ibid.*, p. 230), ou plus généralement le fait que les émotions constituent des complexes motivationnels perturbateurs (*irruptive motivational complexes*). Ainsi, on ne doit pas concevoir la passivité au sens d'inertie, mais plutôt au sens d'une force motivationnelle autonome, qu'il convient de contraster avec des considérations motivationnelles plus « rationnelles ». On ne doit pas non

même catégorie nominale. En effet, une forme de passivité est très certainement à l'œuvre, tant du côté des émotions de base que du côté des émotions cognitives supérieures¹³. La notion de passivité revêt ainsi une certaine valeur, en ce qu'elle permet de regrouper toutes les émotions. Mais la valeur explicative de cette notion est extrêmement limitée. Elle ne rend pas compte, de façon satisfaisante, des principales caractéristiques associées aux émotions de base, comme leur modularité et leur rapidité. Par ailleurs, des propriétés comme la modularité et la rapidité, si elles possèdent une forte valeur explicative dans le cas des émotions de base, ne semblent pas pour autant pouvoir être généralisées à l'ensemble des émotions.

What is to be explained about emotional phenomena in general is the way in which they contrast to other cognitive processes. The phenomena referred to as the "passivity" of emotion are central to this contrast. I am not convinced that all instances of the passivity phenomena can be explained by the modularity of affect programs¹⁴.

À la lumière de la théorie des genres naturels, Griffiths conclut que la cohésion de la sous-catégorie des émotions de base repose sur un mécanisme d'homéostasie causale passablement différent de celui qui semble être à l'œuvre dans le cas des émotions en général. Cette différence s'avère suffisamment importante pour justifier le rejet d'une taxinomie visant à regrouper sous une seule catégorie l'ensemble des émotions.

[We should] refuse to put the higher cognitive emotions in the same category as the affect programs. Whatever psychological mechanism underlies the irruption of these clusters of desires into belief-desire causation, it is not the same mechanism that allows the affect programs to rapidly

plus confondre passivité (le fait que des émotions influent sur la motivation) et modularité (le fait que des émotions ne soient pas aisément modifiées par les croyances).

¹³ *Ibid.*, p. 242.

¹⁴ *Id.*

engage various effector systems without reference to consciously accessible beliefs and desires¹⁵.

Qu'il s'agisse de la passivité des émotions ou de la forme florale des *lilies*, la simple constatation d'une propriété commune ne garantit pas que cette propriété aura la force explicative requise pour unifier la catégorie d'un point de vue épistémique. Au lieu de concevoir les émotions sous l'angle d'une catégorie unique, une taxinomie adéquate proposera plutôt de former deux catégories distinctes correspondant d'une part aux émotions de base et d'autre part aux émotions cognitives supérieures. D'où la nécessité, selon Griffiths, de concevoir en parallèle deux modèles d'étude distincts.

3. Vers l'unité : trois avenues possibles

Une théorie qui souhaiterait malgré tout défendre l'unité de l'ensemble des phénomènes émotionnels devra être en mesure d'exposer et de préciser l'essence commune à l'ensemble des émotions. Trois avenues sont envisageables, mais toutes n'ont pas la même validité.

La première avenue consiste à relever les dénominateurs communs des objets que l'on souhaite unifier sous une même catégorie, puis à élever ces propriétés communes au rang de principe essentiel. Toutefois, comme Griffiths l'a souligné, il est loin d'être garanti qu'une telle tentative soit probante dans le cas des émotions : la passivité, élevée au rang de principe essentiel, ne satisfait pas aux exigences de la théorie des genres naturels. Jesse Prinz, pourtant en faveur d'une thèse unitaire, reconnaît lui-même les écueils d'une telle avenue :

Affect programs and higher cognitive emotions can both be described as irruptive motivation. Iruptive motivation [...] explains why emotions seems passive, drive action, and influence practical reasoning in seemingly irrational ways. [...] I suspect that the irruptive motivation proposal will fail in the end. The problem is that it cast the emotion

¹⁵ *Ibid.*, p. 246.

net too wide. Consider fatigue.[...] Fatigue, it would seem, is an irruptive motivation but not an emotion¹⁶.

Ainsi, cette première avenue n'aurait de validité qu'à deux conditions. D'abord, un principe essentiel doit posséder une forte valeur explicative, comme l'a suggéré Griffiths. Le simple constat d'un dénominateur commun ne suffit pas. Ensuite, il faut éviter l'écueil qui consiste à ratisser trop large, c'est-à-dire à englober des cas qui, selon toute vraisemblance, devraient demeurer hors de la catégorie examinée.

La seconde avenue, qu'on pourrait qualifier de « procustéenne », tente de renforcer l'unité d'une catégorie en excluant, au moyen d'une redéfinition des termes, tous les cas qu'on ne peut ranger sous l'explication théorique privilégiée. L'unité de la catégorie des émotions serait ainsi obtenue en resserrant la catégorie autour des émotions de base — mieux comprises et mieux expliquées —, quitte à exclure de la catégorie les émotions cognitives.

It would be possible to stipulate that instances of anger, disgust, etc., in which the affect programs symptoms are not present are not really instances of those emotions, or are not emotions at all¹⁷.

Dans la mesure où le projet qui nous occupe vise à confirmer ou infirmer l'unité d'une catégorie vernaculaire, cette seconde tentative rate clairement la cible, puisqu'elle propose de fonder l'unité par le biais d'un révisionnisme linguistique. Elle perd de vue l'objectif qui consiste à rendre compte d'une intuition au départ nominale.

La troisième avenue cherche à appréhender les objets moins typiques à la lumière des clés de compréhension qui ont permis d'expliquer les cas les plus évidents. Contrairement à la première avenue, elle ne part pas du recensement de propriétés communes pour aboutir à une essence partagée. Contrairement à la seconde avenue, elle ne cherche pas à exclure d'emblée les cas moins typiques. En fait, cette troisième avenue vise à apparenter les cas moins

¹⁶ Jesse J. PRINZ, *Gut Reactions: A perceptual theory of emotion*, New York, Oxford University Press, 2004, p. 85.

¹⁷ Paul E. GRIFFITHS, *op. cit.*, p. 101.

typiques que sont les émotions cognitives supérieures, aux cas plus typiques, et mieux théorisés, que sont les émotions de base. Certaines de ces stratégies d'assimilation suggéreront que les émotions cognitives supérieures résultent tout simplement du mélange de plusieurs émotions de base, selon un principe de *composition*¹⁸. D'autres suggéreront que les émotions cognitives supérieures consistent en des émotions de base, auxquelles des éléments cognitifs sont surajoutés, selon un principe de *dérivation*. Certains auteurs défendent à la fois la composition et la dérivation¹⁹.

Dans les prochaines sections, j'exposerai la théorie unitaire de Jenefer Robinson et tenterai de voir si cette théorie emprunte l'une ou l'autre des stratégies évoquées ci-dessus. Sans occulter le contraste entre émotions de base et émotions cognitives, Robinson défend, à l'inverse de Griffiths, la pertinence d'une catégorie unique, susceptible de regrouper les deux types de phénomènes, sur la base de leur composante affective.

¹⁸ J'emprunte à la linguistique les termes consacrés *composition* et *dérivation*, qui me semblent particulièrement bien représenter les deux stratégies d'assimilation évoquées. La composition réfère à l'adjonction de deux éléments, par ailleurs autonomes, pour former un troisième élément, comme dans le mot *vide-poche*. La dérivation consiste à joindre un élément autonome (radical) et un élément non autonome (affixe) pour former un troisième élément, comme dans *maisonnette* (l'affixe *-ette* ne s'utilise pas seul). Replacés sur le terrain de la philosophie des émotions, ces termes illustrent, d'une part, la possibilité d'une émotion *composite*, issue du mélange de plusieurs émotions simples qui pourraient, par ailleurs, exister de façon indépendante les unes des autres (horreur = peur + dégoût) et, d'autre part, la possibilité d'une émotion *dérivée*, dont l'élément de base serait, dans un autre contexte, susceptible d'exister indépendamment de l'élément surajouté. En d'autres termes, en considérant les composantes d'une émotion dérivée, on peut concevoir des situations où le radical (émotion de base) n'a pas besoin de l'affixe (élément cognitif), mais l'inverse n'est pas concevable : l'affixe utilisé sans le radical ne fait pas sens.

¹⁹ C'est notamment ce que suggère Jesse PRINZ : « Higher cognitive emotions, one might suppose, are either blends of two basic emotions [...], or combinations of basic emotions and cognitive elaborations. » Jesse J. PRINZ, *op. cit.*, p. 97.

4. Les théories cognitivistes : un repoussoir comportant certaines avancées

Dans l'élaboration de sa thèse, Jenefer Robinson prend soin de bien soupeser les forces et les faiblesses des théories cognitivistes, en particulier le jugementalisme et le quasi-jugementalisme. Bien que ces théories lui servent principalement de repoussoir, Robinson ne les rejette pas pour autant en bloc : elle n'hésite pas, dans certains cas, à intégrer à sa propre position certaines des idées qu'elle juge méritoires, même si celles-ci proviennent précisément des doctrines qu'elle critique.

Le jugementalisme repose sur l'idée qu'une croyance — un jugement, au sens strict — est nécessaire à l'émotion. Un agent ne pourrait pas, par exemple, éprouver une émotion d'amusement, s'il n'entretient pas aussi la croyance que telle ou telle chose est amusante. Dans les versions les plus extrêmes, ce jugement sera également suffisant à l'émotion. Ainsi, la croyance qu'une chose est amusante entraînerait automatiquement l'émotion d'amusement. Le quasi-jugementalisme, pour sa part, propose une version plus souple de la thèse : l'attitude de l'agent n'est plus entendue au sens strict de croyance ou de jugement, mais en un sens beaucoup plus large — on parlera de quasi-jugement —, ce qui permet de mieux rendre compte des cas où l'émotion n'implique pas un véritable engagement de l'agent envers la vérité d'une proposition. Patricia Greenspan résume ainsi les avantages du quasi-jugementalisme par rapport au jugementalisme standard :

In standard cases the evaluative proposition will be one I take to be true. But in order to allow for the possibility that my emotion parts from belief, I shall use "the thought that [the proposition holds]" as a non-committal expression introducing the object of my discomfort, or what I am uncomfortable about²⁰.

Même si, dans les deux cas, l'attitude que l'agent adopte à l'égard d'une proposition n'est pas du même ordre, le jugementalisme et le

²⁰ Patricia S. GREENSPAN, *Emotions and Reasons : an inquiry into emotional justification*, New York, Routledge/Chapman and Hall, 1988, p. 5.

quasi-jugementalisme peuvent tous deux être désignés comme des théories de l'attitude propositionnelle (*propositional attitude theory*). Le jugement ou le quasi-jugement y est considéré comme une condition nécessaire à l'émotion.

Un atout important des théories de l'attitude propositionnelle réside dans l'explication qu'elles fournissent à l'égard de l'*intentionnalité* des émotions. Les émotions portent sur les objets, et il s'agirait, précisément, des objets contenus dans la proposition qui leur est associée²¹. Par ailleurs, les émotions sont dirigées vers ce qui est décisif pour l'agent. Ainsi, suivant l'argument de Robert C. Solomon, les attitudes propositionnelles seraient formées d'une évaluation de ce qui *importe* significativement pour l'agent. "An incident or a perception of an incident alone is never sufficient for emotion, which always involves a *personal evaluation* of the *significance* of that incident."²² Ceci expliquerait, entre autres choses, pourquoi ce ne sont pas toutes les attitudes propositionnelles qui sont aptes à susciter une émotion chez un agent. Enfin, un troisième apport considérable des théories de l'attitude propositionnelle serait leur capacité à *individuer* les émotions à partir des jugements ou des quasi-jugements qui les déterminent. Même si les programmes affectifs semblent pouvoir distinguer physiologiquement certaines familles d'émotion²³, c'est en vertu de la nature du jugement ou du quasi-jugement qu'on pourra adéquatement distinguer les émotions entre elles.

Malgré ces avantages, Robinson croit que les théories cognitivistes devraient être rejetées, parce qu'elles contredisent, de façon flagrante, certaines des données empiriques que nous possédons à propos des émotions. Si les émotions ne sont possibles que dans la mesure où s'y déploie une activité cognitive de type jugemental ou un quasi-jugemental, on ne saurait expliquer la présence des émotions chez les animaux ou chez les nouveaux-nés, qui ne possèdent pas, semble-t-il,

²¹ Griffiths relève toutefois des cas problématiques : définir l'intentionnalité comme propriété essentielle de l'émotion rend difficilement compte du fait que certaines émotions, comme l'angoisse ou l'exaltation, semblent être sans objet. (GRIFFITHS, *op. cit.*, p. 28.)

²² Robert C. SOLOMON *The Passions : emotions and the meaning of life*, Indianapolis, Hackett Publishing, 1993, p. 126.

²³ Jenefer ROBINSON, *Deeper Than Reason : Emotion an its Role in Literature, Music, and Art*, Oxford, Oxford University Press, 2005, p. 33.

les facultés jugées nécessaires à l'émotion. Par ailleurs, Robinson²⁴ cite certaines études de Zajonc qui soulignent que lors d'épisodes émotifs vécus par un agent humain adulte, la manifestation de l'affect peut précéder l'activité cognitive. Zajonc souligne également que le développement de l'affect précède, ontogénétiquement et phylogénétiquement, l'apparition des facultés cognitives supposément nécessaires à l'émotion²⁵. Ces recherches démontrent d'une part que la cognition n'est pas nécessairement ce qui induit l'émotion, du point de vue de la chaîne causale, et d'autre part qu'il existe des épisodes d'émotion sans cognition : ainsi, elles remettent en question à la fois la précédenace temporelle et la nécessité conceptuelle de la cognition.

Un autre problème des théories cognitivistes réside dans leur incapacité à rendre compte des émotions récalcitrantes. Solomon prétend que des changements dans la croyance ou dans le jugement viendront modifier l'émotion en conséquence²⁶. Par exemple, la joie liée à la croyance d'avoir gagné à la loterie s'estompera dès que l'on aura constaté avoir mal relevé les numéros gagnants. Or, les émotions récalcitrantes, comme les phobies, ont ceci de particulier qu'elles semblent complètement imperméables au jugement. On peut croire fermement que les oiseaux ne représentent aucun danger, sans pour autant pouvoir se débarrasser de certaines émotions ornithophobiques. Cette imperméabilité de l'affect à l'égard du jugement est une caractéristique essentielle des émotions, ou à tout le moins, des émotions de base.

Avec ces observations en main, Robinson est en mesure de proposer une théorie non cognitiviste qui soit tout de même capable de rendre compte, grâce à la notion d'évaluation affective (*affective appraisal*), de deux avancées importantes des théories cognitivistes, soit le caractère intentionnel des émotions et leur capacité à évaluer ce qui importe pour l'agent. Par ailleurs, elle confère à la cognition un rôle interprétatif/régulateur (*cognitive monitoring*), qui intervient en fin de processus, et qui permet aux facultés cognitives de réévaluer plus précisément la nature du stimulus de départ; de réajuster, dans la mesure du possible, la réaction affective de l'agent, à la lumière de

²⁴ *Ibid.*, p. 39.

²⁵ Robert B. ZAJONC, « On the Primacy of Affect », *American Psychologist*, vol. 39, no. 2, 1984, p. 117-123.

²⁶ Robert C. SOLOMON, *op. cit.*, p. 126.

cette réévaluation; et, enfin, de discriminer les émotions par le biais d'une réinterprétation conceptuelle. La troisième avancée des théories cognitivistes, l'individuation des émotions, est ainsi sauvegardée.

5. Une structure unique où l'affect prédomine

Ainsi, chez Robinson, tous les phénomènes émotionnels relèveraient de la même structure. L'émotion, en un premier temps, consisterait en une évaluation affective, non cognitive, susceptible de prendre en compte ce qui importe pour l'agent, sans toutefois posséder la précision propre aux jugements cognitifs. En fait, Robinson parle plutôt d'une évaluation rapide, brute, voire grossière, plus proche du réflexe que de la délibération. Dans ce contexte, les pensées et les jugements peuvent être l'objet, au même titre qu'un stimulus externe, d'une évaluation affective. Mais ils ne jouent pas d'autre rôle, à ce stade, que celui de simple stimulus, et ne sont pas, non plus, constitutifs de l'émotion :

The function of non-cognitive affective appraisals is to draw attention automatically and insistently by bodily means to whatever in the environment is of vital importance to me and mine. These affective appraisals can be automatically evoked not only by perceptions such as a sudden loud sound, but also by complex thoughts and beliefs²⁷.

Que les pensées soient parfois aptes à déclencher le processus émotif n'implique nullement qu'elles en fassent partie. L'évaluation affective est précisément ce qui met en marche la deuxième étape du processus, soit une série de réactions physiologiques involontaires, incluant des changements faciaux, une activité du système nerveux autonome et du système moteur, etc. Ces réponses physiologiques auraient pour fonction, entre autres choses, de focaliser l'attention sur certaines caractéristiques saillantes de la situation, afin de préparer l'organisme à l'action. La troisième étape, évoquée précédemment, consiste en une interprétation/régulation d'ordre cognitif. Il convient

²⁷ Jenefer ROBINSON, *op. cit.*, p. 97.

toutefois de noter que, même si cette troisième étape apparaît incontournable dans le cas des émotions d'un sujet humain adulte, il serait probablement incorrect d'en faire une condition nécessaire à l'émotion — à moins, bien sûr, de vouloir soutenir que les animaux et les nouveaux-nés n'éprouvent pas d'émotions véritables, ou encore que leur cognition est suffisamment développée pour permettre de telles réinterprétations. La troisième étape est en quelque sorte facultative.

Quelle est donc, en fin de compte, la place de la cognition dans la théorie des émotions de Robinson? Les jugements ou quasi-jugements ne suffisent pas à déclencher directement une réponse somatique : ils peuvent, tout au plus, agir comme stimuli, et se voir évalués par une faculté affective, qui *elle*, pourra déclencher les réactions somatiques²⁸. L'activité cognitive occupe, dans la théorie de Robinson, un rôle plutôt secondaire : elle précède l'émotion sans en faire partie, à la manière d'un simple stimuli, et la parachève, comme constituant final, mais non nécessaire. Dans le cas d'une émotion cognitive comme la honte, la reconnaissance par l'agent du caractère déshonorant de son état ou de ses actions agirait comme un stimulus interne, de la même façon que la présence de l'ours constitue, dans le cas d'une émotion de base comme la peur, un stimulus externe. En tant que stimuli, ces éléments précèdent l'émotion mais n'en sont pas, à proprement parler, parties intégrantes. Ils en sont plutôt les déclencheurs, au sens où c'est sur eux que portera l'évaluation affective, brute, rapide et imprécise. Et c'est seulement en fin de parcours qu'une véritable évaluation cognitive de la situation pourra permettre de réguler, réinterpréter ou réajuster l'épisode émotif, notamment en corrigeant les imprécisions de la première évaluation.

L'évaluation affective est donc la *première* étape du processus émotif, et s'avère *nécessaire* dans tous les cas, y compris ceux où un jugement cognitif agit en amont du processus émotif. Chez Robinson, la primauté de l'affect s'entend donc à la fois au sens temporel et au sens conceptuel, ce qui n'est pas sans rappeler l'intuition de Phoebe Ellsworth, qui constate que *précédence temporelle* et *nécessité conceptuelle* vont souvent de pair :

²⁸ La complexité du jugement ou du quasi-jugement se voit ainsi « filtrée » par l'évaluation affective, qui n'en retiendra qu'une appréciation basique, du type « Ceci est une menace » ou « Ceci est un danger ». (*Ibid.*, p. 68).

Theorists of emotion [...] differ markedly in the emphasis they place on each of [the] characteristics [of an emotion]. They also differ in their descriptions of the timing of these various components, each author moving his or her favored element to a spot near the beginning²⁹.

6. La réplique de Robinson à Griffiths : la plausibilité de la dérivation

En proposant une explication unique pour l'ensemble des phénomènes émotionnels, basée sur une définition de l'émotion en trois temps où l'affect prédomine — même dans les cas d'émotions cognitives —, Robinson semble répondre aux exigences fixées par Griffiths : une explication causale unique est mobilisée, et celle-ci met en relief les propriétés centrales — et non périphériques — des émotions. Par ailleurs, l'unité de la catégorie ne repose pas sur la notion de passivité contestée par Griffiths.

En outre, il convient de souligner que la rapidité et la modularité des émotions sont conservées par la thèse de Robinson. Paul Griffiths soutenait que ces propriétés étaient hautement typiques des émotions de base, mais impertinentes dans le cas des émotions cognitives supérieures. Les choses semblent différentes sous l'angle de la théorie de Robinson. En observant les cas où des jugements interviennent en amont de l'émotion, comme simples stimuli, on constate que l'évaluation affective, vu son caractère brut et imprécis, ne retient d'eux qu'une appréciation extrêmement sommaire.

I have [...] explained how a primitive, inflexible, non-cognitive appraisal—**This is bad for me!** or **This is a threat!** or **Here's an enemy!**—can take as input a complex cognition, so that I respond with fear not only to

²⁹ Phoebe C. ELLSWORTH, « William James and Emotion : Is a Century of Fame Worth a Century of Misunderstanding? », *Psychological Review*, 1994, vol. 101, no. 2, 1994, p. 227.

the large hairy bear that is haring towards and bearing down on me, but also to my boss's quiet ultimatum³⁰.

Puisque le jugement est ainsi filtré, on peut rendre compte de la modularité de l'émotion. Celle-ci demeure en grande partie imperméable aux détails du jugement.

Qu'en est-il de la rapidité? Comment expliquer la rapidité d'une émotion, lorsque le jugement en amont relève d'une ample et lente délibération? En rejetant le jugement hors du processus émotif, Robinson peut se permettre une réponse simple, qu'elle illustre par l'exemple du rendement des investissements en bourse :

I may feel an emotional response to a realization that my portfolio is performing badly only after having studied long and hard to find out just how the portfolio is behaving. It is only after *lengthy cognitive evaluations* that I realize how badly I am doing, but once I make this discovery, then *all of a sudden* I make an affective appraisal [...] I am suddenly frightened and vulnerable³¹.

C'est donc à partir des émotions de base et de leur structure caractéristique que Robinson définit l'ensemble des émotions. On peut donc penser que sa stratégie s'apparente à la troisième avenue évoquée plus haut, qui consistait à expliquer les cas moins typiques à partir des cas mieux compris : “ It is possible that all emotions can be analysed in terms of basic emotions. ”³² Par ailleurs, il s'agit bien de *dérivation* : l'émotion de base (le radical) pourrait exister indépendamment de la cognition (l'affixe), notamment dans les épisodes qui ne sont ni stimulés ni réinterprétés par l'activité cognitive (chez les animaux ou nouveaux-nés, par exemple). À l'inverse, la cognition (l'affixe) employée indépendamment de toute émotion de base (le radical) ne saurait être qualifiée d'émotion.

³⁰ Jenefer ROBINSON, *op. cit.*, p. 93-94.

³¹ *Ibid.*, p. 62. L'italique est de moi.

³² *Ibid.*, p. 93.

7. Le glissement de l'existential vers l'universel : une critique méthodologique

Malgré la plausibilité des positions défendues par Robinson, un problème subsiste, et il est d'une nature telle qu'il pourrait sensiblement affaiblir la crédibilité de la théorie face aux exigences fixées par Griffiths. En effet, la validité d'une taxinomie des émotions regroupant l'ensemble des émotions sous une catégorie unique — en d'autres mots, la thèse de l'unité — dépend essentiellement de sa capacité à fournir une explication théorique qui ait une solidité dans tous les cas, et à propos de propriétés centrales, non périphériques. Or, Robinson élabore sa thèse à l'encontre des prétentions universalistes des théories de l'attitude propositionnelle : elle cherche à exposer les objections possibles contre l'universalité du modèle jugementaliste ou quasi-jugementaliste. Toutefois, le simple fait d'avoir su mobiliser des arguments contre l'universalité de ces théories ne suffit pas à fonder l'universalité de sa propre théorie.

En logique formelle, la négation d'un quantificateur universel (\forall) appliqué à une proposition peut s'exprimer par l'affirmation du quantificateur existentiel (\exists) appliquée à la proposition niée. Considérons un exemple concret : l'univers du discours étant {les oiseaux}; 'x' représentant le sujet (variable); 'V' représentant la propriété de pouvoir voler. Supposons qu'on souhaite affirmer que tous les oiseaux volent : [$\forall x$] Vx] signifie « pour toutes les valeurs de x dans l'univers du discours, x peut voler ». Si on veut contredire cette affirmation universelle, par exemple en signalant que les autruches ne volent pas, il suffit d'affirmer : [$\exists x$] $\neg Vx$], qui signifie « pour au moins une valeur de x dans l'univers du discours, x ne peut pas voler ». En gros, l'existentialisation d'une proposition peut suffire à invalider l'universalisation de la proposition contraire. Par contre, l'existentialisation d'une proposition ne suffit pas à justifier l'universalisation de cette même proposition.

Zajonc, dans le débat³³ qu'il a entretenu avec Lazarus, a pris soin de signaler les cas où il existe (\exists) une primauté de l'affect sur la

³³ Je réfère principalement ici à deux articles de 1984 : Robert B. ZAJONC, « On the Primacy of Affect », *American Psychologist*, vol. 39, no. 2, 1984, p.

cognition, ce qui lui a permis de démonter aisément les positions de son adversaire, qui affirmait, à l'inverse, que la cognition primait *toujours* (✓) sur l'affect. La contribution de Zajonc dans son article de 1984 se veut modeste mais efficace : elle n'affirme pas une préséance universelle de l'affect, elle se contente de contrecarrer, en citant quelques exemples, les prétentions universalistes de Lazarus.

Il semble qu'en établissant sa position au sein du débat affect/cognition, Robinson soit graduellement passée d'une affirmation existentielle visant à démonter l'universalité de la position cognitiviste, à une affirmation en faveur de l'universalité de l'affect. Ainsi, sans justification suffisante :

Affect *can* precede and therefore does not require cognitive activity³⁴.

An 'appraisal' *can* take place very quickly and prior to any conscious cognition or complex information processing³⁵.

sont rapidement devenus :

At the core of emotion will *always* be physiological responses caused by an automatic affective appraisal³⁶.

An emotion process is *always* triggered by a non-cognitive affective appraisal³⁷.

En d'autres termes, ce qui apparaît initialement comme une tentative d'argumentation contre les théories cognitivistes semble avoir été, un peu trop précipitamment, élevé au rang de règle générale. Pour respecter les conditions fixées par Griffiths, quiconque défend la thèse de l'unité doit pouvoir démontrer qu'il existe une explication satisfaisante en faveur de cette unité, qui soit généralisable à *tous* les cas d'émotion. Cette généralisation doit être fondée; elle ne doit pas relever d'un vice logique. Malgré la plausibilité de certains

117-123; Richard S. LAZARUS, « On the primacy of cognition », *American Psychologist*, vol. 39, no. 2, 1984, p. 124-129.

³⁴ Jenefer ROBINSON, *op. cit.*, p. 39. L'italique est de moi.

³⁵ *Ibid.*, p. 48. L'italique est de moi.

³⁶ *Ibid.*, p. 59. L'italique est de moi.

³⁷ *Ibid.*, p. 77. L'italique est de moi.

aspects de la théorie de Robinson, il semble que son universalisation relève ici d'un défaut méthodologique³⁸. L'ambition de Robinson consiste à démontrer, contre la thèse désunitaire de Griffiths, que *toutes* les émotions appartiennent à la même catégorie : la solidité d'une telle théorie dépend donc, précisément, de son caractère unitaire, c'est-à-dire du fait qu'elle soit universalisable à tous les cas d'émotion. Certes, Robinson vise juste en reprochant aux théories cognitivistes leur incapacité à rendre compte des émotions de base, des phobies ou des émotions récalcitrantes : ainsi, la théorie cognitive n'est donc pas généralisable à tous les cas d'émotion. Mais entre ce simple constat et l'universalité de sa propre théorie, il y a un abîme, que Robinson semble vouloir franchir un peu trop hâtivement, et surtout, sans justification suffisante.

8. Conclusion

À la lumière du conflit entre la thèse désunitaire de Griffiths et la thèse unitaire de Robinson, une question se pose : le conflit peut-il être dissous ? Peut-on tracer une voie mitoyenne entre la position unitaire de Robinson, qui suggère que les émotions puissent être à la fois basiques et cognitives, et la position dualiste de Griffiths, qui propose plutôt que les émotions soient traitées séparément, puisqu'elles appartiennent à l'un *ou* l'autre des deux types, mais jamais à l'un *et* l'autre ?

Je n'espère pas pouvoir développer ici cette voie mitoyenne, du moins, pas au-delà de la simple énonciation d'une hypothèse. Il ne me semble toutefois pas déraisonnable de croire qu'une théorie tripartite puisse être envisagée. En fait, la réfutation des prétentions universalistes de la théorie de Robinson semble mettre à jour la

³⁸ Je ne prétends pas que la théorie de ROBINSON est, dans l'absolu, impossible à universaliser : je cherche plutôt à souligner le fait que la tentative d'universalisation fournie par ROBINSON n'est pas satisfaisante. Par ailleurs, la position de ROBINSON conserve sa plausibilité si on admet d'emblée qu'elle est relative à certaines émotions, et non à toutes les émotions. Mais cette relativisation disqualifie le caractère unitaire de la théorie, et, par voie de conséquence, milite davantage en faveur des thèses de la désunité.

possibilité d'une coexistence des émotions basiques, hybrides, et cognitives. Rien n'empêcherait que des émotions hybrides — cognitives mais dérivées d'émotions de base, à la Robinson — puissent exister aux côtés des émotions non hybrides, c'est-à-dire soit entièrement cognitives soit entièrement basiques. On pourrait ainsi penser que Robinson a raison lorsqu'elle évoque la possibilité que les émotions cognitives soient dérivées des émotions de base, mais qu'elle fait fausse route en prétendant que toutes les émotions se conforment au modèle qu'elle propose. Pareillement, Griffiths aurait raison en affirmant que les émotions peuvent être de nature essentiellement cognitive ou de nature essentiellement basique, mais se méprendrait en récusant l'existence d'émotions à la fois basiques et cognitives. En distinguant deux ensembles (émotions de base, émotions cognitives) joints par une zone commune (émotions hybrides), cette théorie tripartite amenuise l'aspect radicalement dualiste d'une théorie comme celle de Griffiths. Mais il ne s'agit pas, pour autant, de souscrire à une position unitaire, puisqu'on ne peut pas dire qu'une seule explication vaille pour la totalité des phénomènes émotionnels. Il resterait toutefois à voir si cette hypothèse a quelque chance de survie sur le terrain, déjà pleinement occupé, de la philosophie des émotions.

Je n'hésiterai pas toutefois à invoquer, à la lumière de Damasio, la caractère parfois hybridé de la nature elle-même. Plusieurs défenseurs de la thèse de l'unité insistent, dans leur argumentation, sur le fait que des circuits neurologiques normalement associés aux émotions de base sont mobilisés lors de certains épisodes d'émotions dites cognitives. Cette constatation, à leur avis, tend à démontrer que les émotions cognitives sont, d'une certaine façon, assimilables aux émotions de base. Il semble toutefois prématuré de réduire le cognitif à l'affectif sur la base d'un tel constat. Il n'est pas exclu que des processus foncièrement différents puissent emprunter, à un moment ou un autre, la même route :

«La nature, avec son génie du bricolage visant à l'économie, n'a pas élaboré de mécanismes indépendants pour l'expression des émotions primaires et secondaires. Elle a simplement fait en sorte que les émotions

secondaires soient exprimées par le même canal déjà utilisé pour l'expression des émotions primaires³⁹. »

BIBLIOGRAPHIE

- BOYD, Richard, « What realism implies and what it does not », *Dialectica*, vol. 43, no 1-2, p. 5-29.
- DAMASIO, Antonio R., *L'erreur de Descartes*, Paris, Éditions Odile Jacob, 1995, 368 p.
- ELLSWORTH, Phoebe C., « William James and Emotion : Is a Century of Fame Worth a Century of Misunderstanding? », *Psychological Review*, 1994, vol. 101, no. 2, 1994, p. 222-229.
- GOODMAN, Nelson, *Manières de faire des mondes*, Nîmes, Éditions Jacqueline Chambon, 1992, 193 p.
- GREENSPAN, Patricia S., *Emotions and Reasons : an inquiry into emotional justification*, New York, Routledge/Chapman and Hall, 1988, 197 p.
- GRIFFITHS, Paul E., *What Emotions Really Are : The Problem of Psychological Categories*, Chicago, University of Chicago Press, 1997, 286 p.
- JAMES, William, « What is an emotion? », *Collected Essays and Reviews*, New York, Russell & Russell, 1969, p. 244-280.
- KEIL, Frank C., *Concepts, Kinds and Cognitive Development*, Cambridge, MIT Press, 1989, 328 p.
- LAZARUS, Richard S., « On the primacy of cognition », *American Psychologist*, vol. 39, no. 2, 1984, p. 124-129.
- MARTINET, André, *Éléments de linguistique générale*, 3^e éd., Paris, Armand Colin Éditeur, 1991, 221 p.
- PRINZ, Jesse J., *Gut Reactions : A perceptual theory of emotion*, New York, Oxford University Press, 2004, 271 p.
- ROBINSON, Jenefer, *Deeper Than Reason : Emotion an its Role in Literature, Music, and Art*, Oxford, Oxford University Press, 2005, 500 p.
- SOLOMON, Robert C., *The Passions : emotions and the meaning of life*, Indianapolis, Hackett Publishing, 1993, 324 p.

³⁹ Antonio R. DAMASIO, *L'erreur de Descartes*, Paris, Éditions Odile Jacob, 1995, p.183.

ZAJONC, Robert B., « On the Primacy of Affect », *American Psychologist*, vol. 39, no. 2, 1984, p. 117-123.