

Le paradoxe de la fiction : le retour

Florian Cova

Fabrice Teroni

Résumé : Tullmann et Buckwalter (2014) ont récemment soutenu que le paradoxe de la fiction tenait plus de l'illusion que de la réalité. D'après eux, les théories contemporaines des émotions ne fourniraient aucune raison d'adopter une interprétation du terme « existence » qui rende les prémisses du paradoxe incompatibles entre elles. Notre discussion a pour but de contester cette manière de dissoudre le paradoxe de la fiction en montrant qu'il ne prend pas sa source dans les théories contemporaines des émotions. Bien plutôt, son origine se situe dans ce que Radford (1975) décrit comme une incohérence dans nos réactions émotionnelles aux événements fictionnels et non fictionnels. Malgré ce désaccord, nous concédons à Tullmann et Buckwalter qu'une solution satisfaisante au paradoxe de la fiction doit s'appuyer sur les théories des émotions. Ainsi, en guise de conclusion, nous expliquons comment il convient de comprendre l'incohérence en question et comment les théories des émotions pourraient contribuer à la résolution du paradoxe de la fiction.

1. La dissolution du paradoxe de la fiction

S'il peut sembler aujourd'hui évident que les émotions du lecteur jouent un rôle majeur dans la réception et l'appréciation des œuvres de fiction, il est bon de rappeler que cette évidence n'a pas toujours été reconnue. Pendant le XXe siècle, les approches structuralistes et formalistes de l'analyse et de la compréhension des textes littéraires ont en effet proscrit toute référence à l'affectivité du lecteur ou de l'auteur. Les riches travaux de Patrizia Lombardo ont grandement contribué au retour des émotions dans l'étude des textes littéraires et de leur réception.

Toutefois, si l'on considère comme évidente l'idée que l'appréciation des œuvres de fiction est fondamentalement émotionnelle, la manière dont il faut rendre compte de la nature des émotions en question soulève de nombreuses questions philosophiques. Au cœur de ces questions se trouve un paradoxe : le *paradoxe de la fiction*. Depuis sa formulation par Radford (1975), ce paradoxe a engendré une abondante littérature.¹ Il apparaît lorsque l'on considère toutes ensemble les trois propositions suivantes :

(PF1) Nous ressentons de véritables émotions pour les personnages fictionnels et nous soucions de leur sort.

(PF2) Nous ne croyons pas que les entités (personnes, objets, événements) fictionnelles existent, mais avons la croyance que ce qui est décrit dans les œuvres de fiction n'existe pas.

¹ Pour une introduction en français au paradoxe de la fiction, voir Garcia (2014).

(PF3) Nous ne pouvons ressentir de véritables émotions que pour des entités dont nous croyons qu'elles existent réellement.

On remarque aisément le paradoxe que ces propositions semblent engendrer. Si ressentir une véritable émotion requiert de croire que l'entité vers laquelle cette émotion est dirigée existe réellement, et si nous ne possédons pas cette croyance au sujet des entités fictionnelles, alors comment pouvons-nous ressentir des émotions à leur endroit? Etant donné qu'elles semblent incompatibles entre elles, au moins l'une des propositions (PF1), (PF2) ou (PF3) est fautive. *Résoudre* le paradoxe de la fiction consiste à déterminer laquelle.

Une solution qui a rencontré l'assentiment de nombreux philosophes consiste à rejeter (PF1) et à défendre l'idée que nous ne ressentons pas de véritables émotions pour les personnages et autres entités fictionnelles. L'interaction avec les œuvres de fiction produirait en fait des « quasi-émotions », à savoir des états dont la phénoménalité est similaire, voire identique, à celle des véritables émotions (Walton, 1978), mais qui en diffèrent sur un certain nombre d'autres points. Dans cette perspective, il devient tentant de résoudre le paradoxe de la fiction en faisant remarquer que, bien que (PF3) vaille pour les émotions, elle ne vaut pas pour les quasi-émotions. Les véritables émotions requerraient de croire à l'existence de leur objet, alors qu'il suffirait de *faire comme si* un objet existait pour ressentir une quasi-émotion à son endroit. Dans la mesure où il n'est pas absurde de penser que nous interagissons avec les œuvres de fiction en faisant comme si elles décrivaient la réalité², cette solution expliquerait pourquoi ces œuvres suscitent autant de quasi-émotions (Currie, 1990). Quoi qu'on puisse penser de cette solution, ce bref rappel suffit à illustrer l'attitude, jusqu'ici généralement partagée, qui consiste à tenter de *résoudre* le paradoxe de la fiction en identifiant la proposition à rejeter parmi (PF1), (PF2) et (PF3).

C'est à cette attitude générale que Tullmann et Buckwalter (2014) s'opposent dans leur contribution stimulante à ce débat : selon eux, plutôt que de *résoudre* le paradoxe, nous devrions le *dissoudre*. Les trois propositions (PF1), (PF2) et (PF3) seraient en fait compatibles – ce qui revient à dire qu'il n'y aurait, tout compte fait, aucun paradoxe. D'après ces auteurs, si nous avons l'impression que ces propositions sont incompatibles entre elles, c'est seulement parce que nous n'avons pas conscience que le verbe « exister » n'a pas la même signification dans (PF2) et (PF3). Afin de le montrer, Tullmann et Buckwalter

² Nous considérons cette idée brièvement plus bas, voir la section 5.

commencent par distinguer trois sens que le verbe « exister » pourrait avoir en (PF2) : il pourrait signifier (i) être une entité concrète, (ii) pouvoir être une entité concrète, ou (iii) être plus qu'un simple objet imaginaire. Ainsi, la proposition « *nous ne croyons pas que les objets fictionnels existent* » signifierait soit (i) que nous ne croyons pas qu'ils soient des objets concrets, soit (ii) que nous ne croyons pas qu'ils puissent être des objets concrets, ou encore (iii) que nous ne croyons pas qu'ils soient plus que des objets purement imaginaires.

Après avoir distingué ces trois interprétations de (PF2), Tullmann et Buckwalter reportent leur attention sur (PF3). Ils se posent alors la question suivante : avons-nous quelque raison de penser que, dans au moins un des trois sens de « exister » distingués précédemment, ressentir une émotion à propos d'un objet requiert une croyance en son existence? Etant donné que (PF3) porte sur les circonstances et les états mentaux qui donnent naissance à nos émotions, nos auteurs décident de répondre à cette question en passant en revue ce qu'ils tiennent pour les principales théories des émotions. Ils classent ces théories en trois types : les théories faisant appel aux sensations (*feelings*), les théories du jugement et les théories qu'ils décrivent comme purement cognitives. Pour chaque interprétation possible du verbe « exister », ils se demandent alors si les théories en question nous donnent quelque raison d'accepter l'interprétation correspondante de (PF3). Au sortir de leur examen, ils concluent par la négative : aucune de ces théories ne donne à penser que ressentir de véritables émotions requiert de croire en l'existence de leurs objets, et ce pour les trois sens d'« exister » préalablement identifiés. Nous sommes d'accord avec eux sur ce point.³

Tout bien considéré, Tullmann et Buckwalter considèrent qu'il n'y a aucune raison d'accepter une interprétation de (PF3) qui soit en conflit avec (PF2) et que (PF1), (PF2) et (PF3) peuvent donc être réconciliées. Vraisemblablement, l'impression qu'il existe un paradoxe devrait disparaître à la lumière de ces considérations. Et il faut bien admettre que l'approche qu'ils préconisent séduit dans la mesure où elle

³ Ceci ne revient pas à dire qu'aucune théorie concevable des émotions ne pourrait conférer au terme « exister », tel qu'il apparaît en (PF3), l'une des significations qu'il a potentiellement en (PF2), et ne générerait réellement un paradoxe. Tullman et Buckwalter soutiennent simplement et à juste titre (i) que nous n'avons aucune raison d'accepter une telle théorie et (ii) qu'aucune des théories des émotions que nous avons des raisons d'accepter dans les circonstances actuelles ne le fait.

évite la référence (difficilement intelligible) aux quasi-émotions que suppose la solution esquissée plus haut. Mais le paradoxe peut-il vraiment être dissout de cette manière ?

Non, car il existe en fait de bonnes raisons pour résister aux attraits de l'approche défendue par Tullmann et Buckwalter. Dans la section 2, nous expliquons pourquoi la signification d'« exister » dans la proposition (PF3) devrait être fixée à la lumière des phénomènes qui génèrent le paradoxe de la fiction et montrons que ce paradoxe survient parce qu'il y a de bonnes raisons d'adopter une interprétation de (PF3) dans laquelle « exister » signifie « être un objet concret ». Après quoi, dans la section 3, nous expliquons pourquoi la dissolution du paradoxe proposée par Tullmann et Buckwalter se révèle insatisfaisante à la lumière des véritables raisons d'adopter cette interprétation de (PF3). Nous passons ensuite à la partie plus positive de notre discussion et expliquons au cours de la section 4 pourquoi il convient d'aborder d'un point de vue normatif les phénomènes qui fournissent des raisons d'accepter (PF3). À partir de ce nouvel angle d'attaque, nous serons finalement en mesure de consacrer la dernière section à la façon dont les théories des émotions permettent de résoudre le paradoxe de la fiction de façon satisfaisante.

2. Pourquoi la croyance en l'existence serait-elle requise ?

Concentrons-nous, pour commencer, sur la notion de paradoxe. Un paradoxe ne saurait se réduire à un ensemble de propositions incompatibles. C'est pourquoi il ne suffit pas, pour créer un paradoxe, de formuler n'importe quelle triade de propositions incompatibles entre elles telles que (i) la lune est une pastèque géante, (ii) la lune est blanche, et (iii) les pastèques sont vertes. Il faut que les propositions incompatibles soient en outre cognitivement *atrayantes* : (i) ne l'étant pas, l'ensemble constitué des propositions (i), (ii) et (iii) ne crée aucun paradoxe. Il s'ensuit que, si une proposition fait partie d'un paradoxe largement reconnu, il convient, tant que possible, de l'interpréter d'une manière qui la rende cognitivement attrayante. Ceci exige *a fortiori* un examen des raisons qui peuvent être présentées en sa faveur.

La manière dont Tullmann et Buckwalter s'affranchissent de cette exigence eu égard à (PF3) – l'idée que les véritables émotions requièrent une croyance dans l'existence de leurs objets – est étonnante. Pour rappel, leur stratégie consiste à fixer la signification d'« exister » dans cette prémisse à la lumière des principales théories des émotions. Ce qu'il y a d'étonnant dans cette stratégie, c'est que, du propre aveu de

Tullmann et Buckwalter, aucune de ces théories ne plaide en faveur de (PF3). Comme nous l'avons déjà indiqué, nous leur donnons raison sur ce point. Il en résulte néanmoins que nos auteurs ne rendent pas cette prémisse aussi attrayante qu'elle pourrait l'être : son attrait n'a en effet probablement rien à voir avec les théories des émotions.

D'où vient-il alors ? On pourrait affirmer que beaucoup ont été tentés par (PF3) parce qu'ils souscrivaient à une approche non standard des émotions. Etant données les trois familles de théories mentionnées par Tullman et Buckwalter, il sera cependant difficile de désigner une telle approche. D'un autre côté, on pourrait soutenir que (PF3) a séduit parce qu'elle était à tort considérée comme la conséquence d'une théorie standard. Ceci serait néanmoins surprenant, sachant qu'il est assez clair que (PF3) ne résulte pas d'une théorie standard.⁴ Ainsi, il semble qu'aucune raison théorique convaincante ne soit à même d'expliquer l'opinion répandue selon laquelle il y a un paradoxe de la fiction.⁵ Par conséquent, il nous faut rechercher une raison non théorique d'accepter (PF3).

Sur ce point, Tullman et Buckwalter semblent s'en remettre au caractère intuitif de cette proposition. Ils écrivent en effet que la formulation du paradoxe de la fiction par Radford « est composée d'une triade de propositions qui, bien qu'elles se contredisent, sont très intuitives ». (PF3) revêt-elle cependant un caractère intuitif ? Demandez à quelqu'un pour qui le paradoxe de la fiction n'a rien de familier si ressentir une véritable émotion requiert de croire que son objet existe. Pour ce qui est de l'échantillon que nous avons sondé, la réaction typique est l'étonnement : qui a bien pu affirmer une telle chose ? Nos interactions avec les œuvres de fiction semblent constituer autant de contre-exemples à cette thèse, qui se

⁴ Ceci est évident eu égard aux théories qui analysent les émotions en termes de sensations et très clair dans le cas de celles qui les analysent en termes de jugements évaluatifs. En outre, les observations de Tullman et Buckwalter concernant ce qu'ils décrivent comme les théories purement cognitives nous paraissent simples et convaincantes.

⁵ De surcroît, il est difficile de voir à quoi ressemblerait une raison purement théorique d'accepter (PF3). Etant donné le caractère contre-intuitif de cette proposition, une théorie ne l'acceptera que s'il y a des raisons de lier « véritables » émotions et croyances existentielles, et ces raisons sont précisément celles que nous recherchons.

heurte aux intuitions de celui qui, avec la grande majorité des sujets, tient pour acquise (PF1) – l'idée que nous ressentons de véritables émotions pour les personnages de fiction.⁶

Il semble donc que ce ne soit ni en lui cherchant du soutien du côté des théories des émotions ni en insistant sur son caractère intuitif que l'on expliquera l'attrait de (PF3). Il est alors légitime de se demander si cette proposition est vraiment attrayante. En fait, pour s'en convaincre, il suffit de revenir à la formulation originelle du paradoxe par Radford (1975). Ce dernier ne considère pas l'idée que les véritables émotions nécessitent des croyances en l'existence de leur objet comme évidente. Bien au contraire, il consacre une part substantielle de sa discussion à défendre (PF3), ou la proposition légèrement différente employée dans sa propre formulation du paradoxe. L'argument qu'il développe à cette fin prend la forme d'une inférence à la meilleure explication d'un ensemble de phénomènes ordinaires. Les raisons de souscrire à (PF3) ne relèvent donc, selon lui, ni de la théorie des émotions ni de nos intuitions. Pour comprendre ce qui est en jeu ici, rappelons l'exemple qu'il donne :

Imaginez que vous êtes en train de boire un verre avec un homme. Il se met à vous raconter une histoire atroce mettant en scène sa sœur et vous en êtes tout remué. Après avoir savouré votre réaction, il vous avoue qu'il n'a pas de sœur et qu'il a inventé toute l'histoire. Ce qui distingue ce cas des cas précédents, c'est que l'on peut dire que 'l'héroïne' de ce récit est fictionnelle. Néanmoins, [...] il vous est impossible de continuer à être ému une fois l'aveu fait. Certes, il est possible que vous soyez gêné de votre réaction, mais c'est précisément parce qu'elle indique clairement que vous êtes tombé dans un piège – ou alors vous pourrez être embarrassé pour celui qui vous a conté cette histoire et a osé se comporter d'une telle manière. Mais là encore il semble impossible d'être ému sans croire que quelqu'un a souffert. (Radford, 1975, p. 68 ; voir Radford, 2014 pour la traduction française)

Radford offre ici une description fidèle d'une situation assez commune. Il est incontestable que, lorsque nous prenons conscience que les individus sur lesquels portent nos émotions ne sont pas *réels*, c'est-à-dire qu'ils ne sont pas des objets concrets dans le monde actuel, nos émotions tendent généralement à s'effacer, et que nous les percevons rétrospectivement d'un œil critique. C'est cette observation en apparence anodine qui forme la base de l'argument de Radford en faveur de (PF3).

⁶ Deux remarques supplémentaires s'imposent. Premièrement, l'idée que (PF3) n'est pas attrayante ne va pas dans la direction que Tullman et Buckwalter privilégient, puisqu'elle suggère que nous devrions résoudre le paradoxe en rejetant (PF3) plutôt qu'en le dissolvant. Deuxièmement, la stratégie qui consiste à insister sur le caractère intuitif de (PF3) semble également être en tension avec leur tentative de dissoudre le paradoxe de la fiction. Nous n'insisterons pas sur ce problème ici.

Son argument s'appuie en effet sur un certain nombre d'observations « anthropologiques » qui mettent en lumière un aspect somme toute assez banal de notre vie affective. Si vous êtes attristé(e) par les souffrances présumées de la tante de votre voisin, que vous croyez séquestrée dans un pays lointain, votre émotion cesse, plus ou moins abruptement, lorsque vous apprenez que votre voisin n'a jamais eu de tante. De la même manière, si vous vous inquiétez pour un ami, que vous croyez en danger, votre peur s'estompe sitôt que vous apprenez qu'il est en sécurité. Pourquoi les émotions sont-elles susceptibles de tels changements? Radford propose une explication pleine de bon sens : pour être attristé par les circonstances qui affectent une personne, il faut croire que cette personne souffre, et donc qu'elle existe. Si cette condition cesse d'être satisfaite, l'émotion s'évanouit ou, si elle ne le fait pas, est tenue pour « injustifiée » ou « inappropriée » – de la même manière que l'on évaluerait comme inappropriée votre peur récalcitrante d'une araignée que vous savez par ailleurs totalement inoffensive.

Toutefois, cette description de notre vie affective s'accorde difficilement avec les réactions émotionnelles que suscite la fiction, dans la mesure où ces dernières subsistent presque toujours en dépit du fait que nous ne croyons pas en l'existence de leur objet. Ceci pourrait nous inciter à rejeter (PF3), mais nous serions alors dans l'impossibilité d'expliquer les attitudes qu'illustrent les cas présentés par Radford. Nous rencontrons ainsi une manière de dilemme. D'un côté, nous pouvons rejeter (PF3), mais alors nous ne parviendrons pas à expliquer le fait que nos émotions s'évanouissent souvent lorsque nous découvrons que leur objet n'existe pas. De l'autre, nous pouvons maintenir (PF3), mais alors nous ferons de nos réactions émotionnelles à la fiction un phénomène déconcertant et incohérent, ce qui semble tout aussi insatisfaisant.

Ainsi, Radford ne dit pas que nous devrions adopter (PF3) sur la seule base de son caractère intuitif ou d'une théorie particulière, mais parce qu'elle fournit la meilleure explication d'une série de phénomènes. Le paradoxe de la fiction lui semble être enraciné dans cette tension entre, d'une part, ce qui paraît constituer la meilleure explication de notre vie affective dans tout une gamme de cas ordinaires et, d'autre part, la certitude introspective que nous avons de ressentir de véritables émotions pour des entités fictionnelles dont nous savons qu'elles n'existent pas réellement. C'est pourquoi Radford décrit le phénomène qui l'intéresse comme une sorte d'« incohérence » entre nos réactions émotionnelles face aux entités fictionnelles, d'une part, et nos réactions émotionnelles face aux entités non fictionnelles, d'autre

part. Résoudre le paradoxe de la fiction implique de rendre compte de cette incohérence – que nous appellerons *l'incobérence affective*.

Si c'est là que le paradoxe de la fiction trouve sa source, alors l'interprétation d'« exister » que nous devons adopter en le formulant est celle qui permettra d'expliquer l'ensemble de phénomènes en question. Or, (PF3) ne peut contribuer à cette explication que si « exister » est compris comme signifiant « être un objet concret » ou « être un habitant du monde actuel ». Et, puisque nous ne croyons évidemment pas que les personnages de fiction habitent notre monde comme autant d'entités concrètes que nous pourrions croiser au coin de la rue, il semble finalement qu'« exister » ait la même signification en (PF2) qu'en (PF3), et donc que le paradoxe de la fiction subsiste malgré les tentatives pour le dissoudre.

3. Les théories contemporaines peuvent-elles dissoudre le paradoxe ?

Nous avons conclu la section précédente par ce constat : les raisons qui ont été avancées en faveur de (PF3) nous imposent d'interpréter « exister » uniformément dans les trois prémisses qui forment le paradoxe de la fiction. Mais alors, la solution de Tullman et Buckwalter qui consistait à dissoudre ce paradoxe en faisant valoir qu'aucune théorie des émotions ne permet l'interprétation requise de (PF3) est-elle encore à portée de main ?

Voici une première raison d'en douter. A la lumière de ce qui précède, les théories des émotions ne permettent pas de dissoudre le paradoxe – elles n'impliquent pas que les propositions (PF1), (PF2) et (PF3) puissent être conciliées. Ces théories nous invitent plutôt à *résoudre* le paradoxe en rejetant l'une de ses prémisses, en l'occurrence (PF3), ce qui est nettement plus conforme à la tradition.

Deuxièmement, le fait qu'elle ne bénéficie pas du soutien des théories des émotions ne justifie pas que l'on rejette (PF3). Une telle approche serait par trop superficielle, dans la mesure où elle ne rendrait pas compte des phénomènes qui ont amené Radford à formuler le paradoxe. En l'adoptant, on renoncerait à expliquer pourquoi il n'est pas requis de croire en l'existence d'Anna Karénine pour être attristé(e) par sa situation, alors que le même type de croyance semble jouer un rôle essentiel dans la plupart des contextes non fictionnels. Radford lui-même envisage, avant de la rejeter, cette approche qui revient à considérer les réactions émotionnelles que suscite en nous la fiction comme autant de données psychologiques brutes. Ce qu'il juge insatisfaisant :

Le problème qui nous occupe est que les gens puissent être émus par les souffrances fictionnelles malgré le fait brut que, dans d'autres contextes, la croyance dans la réalité d'une souffrance à laquelle ils assistent ou qu'on leur a rapportée est nécessaire pour qu'ils aient cette même réaction. (Radford, 1975, p. 72 ; voir Radford, 2014 pour la traduction française)

Comme Radford, nous considérons cette solution comme une manœuvre de dernier recours, puisqu'elle ne soutient pas la comparaison avec les autres solutions disponibles. Considérons deux solutions parmi d'autres : (i) lorsque nous sommes immergés dans la fiction, nous croyons, ou au moins nous acceptons, que les personnages et les événements fictionnels existent (par exemple : Suits, 2006 ; Todd, 2013), (ii) les états affectifs que nous ressentons pour des individus ou des événements dont nous ne croyons pas qu'ils existent ne sont que des quasi-émotions (par exemple : Walton, 1978 ; Currie, 1990). En dépit des difficultés qu'elles peuvent soulever, ces solutions ont au moins le mérite de dépasser le stade du constat pour s'intéresser aux phénomènes sous-jacents. La première réconcilie les propositions (PF1) et (PF3) : elle implique à la fois que nous ressentons de véritables émotions à l'égard des entités que nous considérons être fictionnelles et que les véritables émotions requièrent des croyances existentielles. La seconde nous permet de rendre compte, d'une part, des phénomènes illustrés par Radford, parce qu'elle souscrit à (PF3) et, d'autre part, de l'impression que nous avons de ressentir de véritables émotions face à la fiction.⁷ La stratégie privilégiée par Tullmann et Buckwalter n'est pas faite du même bois, puisqu'elle ne prétend pas expliquer l'incohérence affective. Comme cette incohérence est à la source du paradoxe de la fiction, ces auteurs échouent donc tant à le résoudre qu'à le dissoudre, et ce même en admettant que (PF3) va à l'encontre des théories des émotions actuelles.

Une troisième raison de rejeter la stratégie de Tullman et Buckwalter est qu'une version de (PF3) à même de générer le paradoxe pourrait bien ne pas s'opposer à ces théories. Bien sûr, la version que nous avons employée jusqu'ici (c'est-à-dire : « Nous ne pouvons ressentir de véritables émotions que pour des entités dont nous croyons qu'elles existent réellement ») est tout autant inconciliable avec ces théories qu'avec les données de l'introspection. Après tout, la seule idée de manger de la glace parfumée au vomit suscite chez la plupart d'entre nous le dégoût, sans pour autant que nous croyions en l'existence d'un tel mets. Mais le paradoxe ne requiert pas une version si forte de (PF3) : une version plus faible ferait tout

⁷ Compte tenu du degré de ressemblance entre les quasi-émotions et les émotions véritables, il n'y en effet rien d'étonnant à ce que nous soyons susceptibles de les confondre.

aussi bien l'affaire, selon laquelle un sous-ensemble seulement des émotions que nous ressentons au sujet d'entités fictionnelles présuppose une croyance en leur existence. Radford ne s'opposerait pas à cette version de (PF3). En effet, il n'affirme nulle part que *toutes* les émotions requièrent une croyance en l'existence de leur objet. Il soutient seulement que c'est le cas de la tristesse, tout en remarquant que les œuvres de fiction la provoquent parfois. Cette version affaiblie de (PF3) ne s'oppose pas franchement aux théories standard des émotions : comme Tullmann et Buckwalter le remarquent, ces théories impliquent seulement que certaines émotions peuvent être provoquées en l'absence de croyances en l'existence de leur objet, ce qui n'exclut pas que d'autres émotions puissent nécessiter de telles croyances pour être suscitées.

En somme, tout porte à croire que l'on ne peut ni résoudre ni dissoudre le paradoxe de la fiction simplement en affirmant qu'une version forte de (PF3) s'oppose aux théories standard des émotions.

4. Approches descriptives et normatives de l'incohérence affective

Malgré ces critiques, nous rejoignons Tullmann et Buckwalter sur leur tentative d'aborder le paradoxe de la fiction à la lumière des récents développements en théorie des émotions, et nous partageons leur conviction que (PF3) en constitue le maillon faible. Cependant, afin de dépasser ce que nous avons diagnostiqué dans la section précédente comme une forme d'indifférence à l'égard des phénomènes qui motivent l'adoption de cette proposition, il convient de solliciter l'approche théorique des émotions d'une manière plus substantielle. Cette dernière doit en effet nous fournir une explication de ce que nous avons appelé « l'incohérence affective », c'est-à-dire la tension apparente entre la manière dont les émotions se comportent dans les cas fictionnels et dans les cas non fictionnels. Malheureusement, s'il est légitime d'attendre cela d'une solution satisfaisante, c'est en vain que l'on chercherait une telle explication dans la littérature. Rien n'indique pour autant que nous ne puissions pas nous inspirer de discussions récentes. Dans cette section, nous distinguons deux façons de rendre compte de l'incohérence affective – l'une situe cette incohérence au niveau descriptif, l'autre au niveau normatif – et nous plaidons pour la seconde. Ce qui nous conduira à examiner ensuite les explications normatives de cette incohérence.

Commençons par une affirmation évidente : la différence dans la manière dont les émotions se comportent dans les cas fictionnels et dans les cas non fictionnels se laisse décrire comme une *incohérence au*

niveau de l'impact cognitif de l'information. Les émotions sont cohérentes dans le sens pertinent lorsqu'elles font preuve de la même sensibilité à la même information dans les situations tant fictionnelles que non fictionnelles. Comme nous l'avons vu, les émotions que nous ressentons sont bien souvent incohérentes : bien qu'elles soient sensibles à l'information à propos de l'existence de leur objet dans les cas ordinaires, elles ne le sont pas dans les cas fictionnels. Ceci étant dit, il est crucial de comprendre que nous pouvons nous attaquer à l'incohérence en question à deux niveaux distincts – l'un purement *descriptif*, l'autre *normatif*.

La stratégie qui consiste à s'en tenir au niveau *descriptif* est insatisfaisante, car elle fait face au dilemme suivant. On pourrait soutenir que l'incohérence n'est qu'apparente, mais alors on échouerait à rendre compte des phénomènes illustrés par les exemples de Radford. C'est la première branche du dilemme. Tant que nous nous cantonnons au niveau descriptif, la seule alternative consiste à considérer l'incohérence affective comme un fait brut. On aura reconnu l'approche superficielle et insatisfaisante anticipée par Radford, et vers laquelle Tullmann et Buckwalter semblent contraints de se tourner. C'est la seconde branche du dilemme.

Ceci suggère fortement que l'incohérence affective devrait être abordée au niveau *normatif* plutôt que descriptif. Au niveau normatif, la question pertinente est la suivante : y a-t-il quelque chose qui cloche dans le fait que nos émotions répondent à l'information différemment selon que leurs objets sont fictionnels ou non fictionnels ? Ces réactions sont-elles inappropriées, incorrectes ou irrationnelles ? Supposons que la rationalité des émotions qui portent sur des entités fictionnelles ne soit pas exactement calquée sur la rationalité des émotions qui portent sur des entités non fictionnelles, de telle sorte que ce qui apparaît comme incohérent d'un point de vue descriptif soit cohérent d'un point de vue normatif. Cela signifierait qu'il n'y a, tout compte fait, pas de dilemme : nous serions en position de reconnaître l'incohérence affective tout en l'expliquant au niveau normatif. Nous allons maintenant introduire deux thèses à propos du caractère approprié des émotions, dont nous pensons qu'elles confirment cette idée – ces thèses nous aideront à comprendre pourquoi une incohérence affective au niveau descriptif constitue une pré-condition pour que les émotions qui portent sur des entités fictionnelles puissent être rationnelles.

Il n'est pas aisé de répondre à la question : « sous quelles conditions une réaction émotionnelle est-elle rationnelle, appropriée ou correcte ? » Néanmoins, dans le cadre de la présente discussion, il suffira de

souligner deux idées qui sont au centre des réponses les plus récentes.⁸ Selon la première, les émotions ont des conditions de correction évaluatives. Il est correct ou approprié d'être triste à propos d'un événement donné dans la mesure où cet événement constitue une perte ; de la même manière, il est correct ou approprié d'être en colère à l'égard d'une personne dans la mesure où elle s'est comportée de manière offensante. La seconde idée est que les émotions dépendent d'autres états mentaux qui fonctionnent, pour ainsi dire, comme leurs « bases cognitives ». Pour être attristé par un événement, vous devez vous le représenter d'une manière ou d'une autre (peut-être l'avez-vous vu, ou avez-vous inféré à l'instant une croyance à son propos).

Une manière attrayante de combiner ces deux idées consiste à dire que les émotions héritent une partie de leurs conditions de correction de ces bases cognitives, auxquelles elles ajoutent une couche évaluative (Deonna et Teroni, 2012 : chap. 7). Supposons qu'il vous semble que vous voyiez telles et telles propriétés d'un événement donné. Cette expérience est correcte si, et seulement si, l'événement en question exemplifie les propriétés que vous semblez voir. Si maintenant vous réagissez avec tristesse à ce que vous voyez, votre tristesse est correcte si, et seulement si, l'événement que vous voyez constitue une perte. C'est pourquoi le caractère approprié des émotions devrait être discuté à la lumière, premièrement, du caractère approprié de leurs bases cognitives et, deuxièmement, des propriétés évaluatives qu'exemplifient leurs objets.

Revenons-en à l'incohérence affective considérée au niveau normatif, et posons-nous plus spécifiquement la question suivante : le fait que les émotions qui portent sur des entités fictionnelles se comportent différemment de celles qui portent sur des entités non fictionnelles signifie-t-il que les premières sont inappropriées ou irrationnelles ? Selon que l'on rangera le problème du côté des *propriétés évaluatives* ou de celui des *bases cognitives*, il y a deux raisons potentielles de répondre à cette question par l'affirmative, et ainsi de conclure que l'incohérence affective est problématique au niveau normatif.

Pour ce qui est des *propriétés évaluatives*, on pourrait affirmer que les entités fictionnelles ne sont pas à même de les exemplifier. On pourrait ajouter que la manière dont nous réagissons à l'information dans les

⁸ Pour une discussion et des références à la riche littérature consacrée aux questions sous-jacentes, voir Teroni (2014).

cas non fictionnels semblables à ceux présentés par Radford manifeste notre compréhension intuitive du fait que les entités non existantes ne peuvent pas exemplifier de propriétés évaluatives, alors que cette compréhension n'a pas de prise sur nos réactions émotionnelles dans les cas fictionnels.⁹ Pour cette raison, l'incohérence affective infecterait le niveau normatif. Mais ce raisonnement n'est pas convaincant. Premièrement, il est fondée sur une assertion ontologique gratuite – même en admettant qu'Ivanhoé ne puisse pas faire face au danger et, de la même manière, que l'Oncle Picsou ne puisse pas être insulté, nous aimerions savoir pourquoi. Deuxièmement, en l'absence d'un argument, il est difficile d'admettre qu'une forme de compréhension que nous déployons aisément dans d'autres contextes échoue systématiquement à imprégner nos réactions affectives à la fiction – on a du mal à croire qu'il s'agisse là d'un fait brut.

Concentrons-nous maintenant sur l'autre alternative, qui consiste à ranger le problème du côté des *bases cognitives*. On pourrait suggérer que la base cognitive d'une émotion à propos d'une entité fictionnelle n'est autre que la croyance que cette entité existe actuellement et, par là, rejeter (PF2). Suite à quoi on argumenterait que l'incohérence affective demeure problématique au niveau normatif, de la manière suivante. Les normes qui portent sur les croyances sont suffisamment stables à travers les contextes pour que l'information qu'un objet n'existe pas, si elle rend approprié d'abandonner une croyance dans un contexte non fictionnel, ait le même effet sur la même croyance dans un contexte fictionnel. Il s'agit d'une stratégie possible, mais qui conclut que l'incohérence affective est normativement problématique pour la seule raison qu'elle dépeint – sans justification apparente – les consommateurs de fiction comme de grands naïfs : les émotions qu'ils ressentent seraient basées sur des croyances fausses et pour le moins bizarres.

En l'absence de raisons plus convaincantes d'affirmer que l'incohérence affective infecte le niveau normatif, peut-être faut-il admettre que la rationalité des émotions dont les objets sont fictionnels n'est pas strictement équivalente à la rationalité des émotions dont les objets ne le sont pas. Conserver une émotion à propos d'une entité fictionnelle tout en étant informé qu'elle n'existe pas (i.e. qu'elle n'est pas un objet concret dans le monde réel) pourrait bien être rationnel. Reste à expliquer l'incohérence affective au niveau normatif. Dans la dernière section, nous examinons deux tentatives qui vont dans ce sens.

⁹ A moins, bien sûr, que ces réactions ne constituent pas de véritables émotions.

5. Deux explications normatives de l'incohérence affective

Une manière naturelle d'aborder l'incohérence affective au niveau normatif revient à se pencher sur la nature des bases cognitives des émotions. Comme nous allons le voir, cela donne lieu à deux approches intéressantes, qui ont en commun de considérer que cette incohérence n'est pas normativement problématique en distinguant les bases cognitives des émotions qui portent sur la fiction des bases cognitives des autres émotions. Mais, tandis que la première trace la distinction au niveau des *modes psychologiques* qu'elles renferment, la seconde le fait au niveau de leurs *contenus*.

Selon la première approche, s'il est approprié de conserver une émotion dont l'objet est fictionnel tout en sachant que cet objet n'existe pas, c'est parce que les bases cognitives des émotions dont les objets sont fictionnels ne sont pas des croyances, mais des épisodes de supposition, d'imagination ou d'« acceptation ». Cette approche est attrayante dans la mesure où, comme nous l'avons souligné, les bases cognitives des émotions sont susceptibles de varier considérablement. Ainsi, simplement imaginer un objet dégoûtant peut suffire à susciter le dégoût, et ce en l'absence de la croyance que cet objet existe. Pour autant, le dégoût est parfois provoqué par d'autres modes psychologiques, comme la croyance. C'est par exemple le cas si nous découvrons qu'un cafard mort se trouve dans la boisson que nous sommes sur le point d'ingérer. Bien sûr, le fait que des émotions d'un même type puissent avoir différentes bases cognitives a un impact sur le genre d'information auquel elles seront sensibles : si le dégoût est sensible au changement de croyance lorsqu'il a pour base cognitive une croyance, il ne l'est pas quand sa base cognitive est un épisode d'imagination (Rozin, Milliman & Nemeroff, 1986).

La présente approche s'inspire de ces observations. Elle soutient que, contrairement aux émotions dont les objets ne sont pas fictionnels, qui sont (typiquement) basées sur des croyances, les émotions qui portent sur des entités fictionnelles sont basées sur des états mentaux dont le mode psychologique n'est pas « sérieux » – c'est-à-dire des états dont le contenu est imaginé, supposé ou accepté sans pour autant que le sujet s'engage envers sa vérité. Croyances et suppositions étant sujettes à des normes différentes, les émotions doivent réagir à l'information de manières différentes, selon qu'elles portent sur des entités fictionnelles ou non fictionnelles. Il existe, par exemple, une norme selon laquelle nous devrions abandonner la croyance que p à la lumière de l'information que $non-p$, mais il est évident qu'aucune norme de la sorte ne porte sur la supposition, l'imagination ou l'acceptation. Le fait que les émotions répondent

différemment à l'information selon qu'elles portent sur des entités fictionnelles ou non fictionnelles serait une conséquence du fait qu'elles sont basées sur des modes psychologiques non sérieux. En outre, qu'une émotion dont l'objet est fictionnel soit basée sur une supposition ne signifie pas qu'elle soit inappropriée ou irrationnelle. Cela indique au contraire que celui qui la ressent n'est pas naïf. Nous tenons donc une première explication normative de l'incohérence affective: les émotions dont les objets sont fictionnels étant basées sur des états mentaux dont le mode psychologique n'est pas sérieux, il est approprié qu'elles demeurent, au contraire des émotions dont les objets sont non fictionnels, insensibles à l'information que leur objet n'existe pas. Il s'agit même d'une pré-condition de leur rationalité.

Aussi attrayante qu'elle puisse paraître, cette première approche soulève les difficultés suivantes. Premièrement, il n'est pas si clair que la supposition ou l'imagination constitue le mode psychologique typique de notre engagement envers la fiction. Comme certains auteurs l'ont souligné à juste titre, il ne semble pas y avoir quoi que ce soit « de supposé, d'hypothétique ou de provisoire » dans nos pensées à propos d'Anna Karénine (Matravers, 1991 ; voir aussi Neill, 1993). Il semble inexact de dire que « nous faisons comme si » (nous croyions que) la vie d'Anna avait un dénouement tragique. Ces descriptions sont pourtant correctes quand nous songeons à des cas d'imagination ou de supposition. Deuxièmement, souligner le contraste entre croire et faire comme si n'est probablement pas la meilleure option, puisque cette distinction s'applique également à notre engagement envers les entités réelles et fictionnelles. On peut croire que *on* faire comme si son mari fait face à une situation tragique, de même que l'on peut croire que *on* faire comme si Anna Karénine fait face à une situation tragique – et il est plausible d'affirmer que cette différence dans les modes impliqués affecte les émotions qui s'ensuivent. Par conséquent, pourquoi insister sur le fait que notre engagement envers la fiction se construit exclusivement sur des modes psychologiques non sérieux ? Puisque nous semblons entretenir de nombreuses croyances à propos de la fiction, pourquoi ne pas construire une approche autour d'elles ?

Ceci nous amène à la seconde approche, selon laquelle, s'il est approprié de conserver des émotions dont les objets sont fictionnels en dépit de notre croyance que ces objets n'existent pas, c'est en vertu du contenu, et non du mode psychologique, de leurs bases cognitives. Selon les partisans de cette approche (par exemple : Matravers, 1991 ; Neill, 1993), les émotions en question ne sont pas basées sur des épisodes de supposition, mais sur des croyances. Seulement, ces croyances ont un contenu plus complexe que celles

qui sont responsables des émotions dont les objets ne sont pas fictionnels : leur contenu est préfacé d'un opérateur de fiction. Ainsi, nous croyons que, *dans la fiction*, la vie d'Anna a un dénouement tragique. Ceci explique pourquoi les croyances pertinentes réagissent, et pourquoi elles doivent réagir, de manières différentes à l'information. Le type d'information auquel une croyance doit être sensible dépend évidemment de son contenu : la croyance que, *dans la fiction*, Anna Karénine commet un suicide ne devrait pas être révisée à la lumière de l'information qu'elle n'existe pas dans le monde actuel. L'information à laquelle cette croyance doit être sensible concerne ce qui arrive à Anna *dans le monde fictionnel*, et elle est exclusivement issue du roman de Tolstoï (voir sur ce point : Livingston & Mele, 1997). Ceci génère une seconde explication normative de l'incohérence affective: les émotions dont l'objet est fictionnel étant basées sur des croyances dont le contenu est préfacé par un opérateur de fiction, il est approprié qu'elles demeurent, au contraire des émotions dont les objets sont non fictionnels, insensibles à l'information que leur objet n'existe pas.

L'avantage principal de cette seconde approche est qu'elle évite de postuler le genre d'attitudes non sérieuses ou « hypothétiques » vis-à-vis des objets fictionnels qu'il semble difficile de concilier avec la nature de notre engagement envers la fiction. Mais cette approche soulève un problème important : elle pourrait bien être incompatible avec le genre d'« immersion » qui est, selon certains auteurs, caractéristique de cet engagement (cf. Todd, 2013). La première approche avait l'avantage d'expliquer facilement ce phénomène. Lorsque nous imaginons que *p*, notre attention est typiquement focalisée sur ce que nous imaginons plutôt que sur le fait que nous l'imaginons ; c'est pourquoi nous sommes si souvent embarqués ou immergés dans la fiction. L'introduction d'un opérateur de fiction dans le contenu des modes psychologiques en jeu dans notre engagement envers la fiction ne facilite pas la tâche d'expliquer pourquoi il en va ainsi. Compte tenu de la nature de ce contenu, il serait impossible de se concentrer sur ce que l'on croit – que, *dans la fiction*, Anna Karénine commet un suicide – sans penser qu'« il ne s'agit que d'une fiction ». Or, l'immersion suppose au contraire que nous l'omettions.

Les deux approches que nous venons d'esquisser soulèvent bien sûr un certain nombre de problèmes, mais elles illustrent la manière dont les théories contemporaines pourraient être développées afin de rendre justice aux émotions dont les objets sont fictionnels et de résoudre le paradoxe de la fiction. On pourra certes trouver ces réflexions très embryonnaires, mais notre but n'était pas de résoudre le

paradoxe de la fiction. Il était d'expliquer pourquoi on ne peut pas le *dissoudre* au moyen des théories contemporaines des émotions. Le paradoxe résiste à ce type de tentatives parce qu'il articule une tension psychologique profonde : pourquoi les émotions qui ont pour objets des entités fictionnelles se comportent-elles si différemment des émotions qui portent sur d'autres entités ? Dans la mesure où cette tension constitue la source du paradoxe, une solution ou une dissolution attrayante ne saurait se contenter de l'ignorer.

Bibliographie

- Currie, G. (1990). *The Nature of Fiction*. Cambridge: Cambridge University Press.
- Deonna, J. et Teroni, F. (2012). *The Emotions: A Philosophical Introduction*. New York: Routledge.
- Garcia, A. (2014). Anna Karénine et moi. *REPHA* 7, 81-89.
- Livingston, P. & Mele, A. (1997). Evaluating Emotional Responses to Fiction. Dans M. Hjort & S. Laver (dir.), *Emotions and the Arts* (pp. 157-176). New York: Oxford University Press.
- Matravers, D. (1991). Who's Afraid of Virginia Woolf? *Ratio* 4, 25-37.
- Neill, A. (1993). Fiction and the Emotions. *American Philosophical Quarterly* 30, 1-13.
- Radford, C. (1975). How Can we Be Moved by the Fate of Anna Karenina? *Proceedings of the Aristotelian Society Supplementary Volume* 49, 67-80.
- Radford, C. (2014). Comment pouvons-nous être émus par le sort d'Anna Karénine ? (trad. par F. Cova et A. Garcia) *REPHA* 7, 91-107.
- Rozin, P., Millman, L. & Nemeroff, C. (1986). Operations of the Laws of Sympathetic Magic in Disgust and Other Domains. *Journal of Personality and Social Psychology* 50, 703-712.
- Suits, D. (2006). Really Believing in Fiction. *Pacific Philosophical Quarterly* 87, 369-386.
- Teroni, F. (2014). Emotions et connaissance. Dans J.-M. Chevalier et B. Gaultier (dir.), *Connaître: Questions d'épistémologie contemporaine* (pp. 9-34). Paris: Ithaque.
- Todd, C. (2013). Attending Emotionally to Fiction. *Journal of Value Inquiry* 46, 449-465.
- Tullmann, K., & Buckwalter, W. (2014). Does the Paradox of Fiction Exist? *Erkenntnis* 79, 779-796.
- Walton, K. (1978). Fearing Fictions. *Journal of Philosophy* 75, 5-27.