



HAL
open science

Philosophie de l'esprit et psychologie, du transcendantal à l'empirique

Pascal Engel

► **To cite this version:**

Pascal Engel. Philosophie de l'esprit et psychologie, du transcendantal à l'empirique. *Psychologie Française*, 1999, 44 (3), pp.257- 264. <ijn_00000318>

HAL Id: ijn_00000318

https://hal.science/ijn_00000318v1

Submitted on 26 Jan 2003

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

PHILOSOPHIE DE L'ESPRIT ET PSYCHOLOGIE : DU TRANSCENDANTAL A L'EMPIRIQUE

Pascal Engel

Psychologie française, 44, 3, sept 1998, 257-264

On dit souvent que les progrès des sciences cognitives ont créé les conditions d'un remariage entre la philosophie et la psychologie, un siècle après leur retentissant divorce. Si l'on en juge par l'abondance, depuis au moins une vingtaine d'années, de la production en « philosophie de l'esprit », en « philosophie cognitive » ou « philosophie de la psychologie »^[1], et l'existence, en langue anglaise, de revues comme *Mind and Language* et *Philosophical Psychology*, ou d'associations comme l'*European Society for Philosophy and Psychology*, ainsi que par l'existence (encore rare en France) d'institutions et de diplômes interdisciplinaires en sciences cognitives accueillant des philosophes en leur sein, il y a lieu d'être assez optimiste. Mais aussi optimiste soit-on sur ces plans, c'est à un pessimisme plus profond et plus durable qu'il faut répondre. Il peut, très naturellement, venir de chaque camp. Du côté des psychologues, on peut se demander, à supposer qu'il soit nécessaire de prendre en compte certaines analyses philosophiques portant sur l'esprit, si elles contribueront réellement à l'obtention de vérités *sur la cognition humaine*. Du côté des philosophes on peut se demander, à supposer qu'il soit vraiment nécessaire de prendre en compte des vérités quant au fonctionnement et à la structure de la cognition, si les vérités en question seront réellement des vérités *philosophiques*.

On peut, me semble-t-il, partager ce double pessimisme, sans pour autant défendre l'idée que *toutes* les vérités philosophiques sont nécessairement des vérités conceptuelles ou *a priori*, soustraites par définition à tout contrôle expérimental, et sans, inversement, soutenir, comme le font nombre de philosophes « naturalistes », qu'il n'existe en fait qu'une différence de degré entre les théories scientifiques et les théories philosophiques, et donc pas de distinction de principe entre des vérités de type empirique et des vérités de type conceptuel ou *a priori*. En d'autres termes, je voudrais soutenir que bien qu'il y ait, en philosophie de l'esprit, un certain nombre de thèses et de concepts qui ne peuvent pas être soumis à des tests empiriques sur la

cognition, il existe une zone intermédiaire de concepts qui se situent à l'interface entre philosophie et sciences cognitives et qui permettent de répondre de manière raisonnablement optimiste aux questions précédentes. J'appellerai cette zone intermédiaire, pour reprendre un terme kantien, « psychologie transcendantale », et je voudrais suggérer que cette psychologie en apparence assez suspecte ne devrait pas rebuter outre mesure les psychologues. Il m'est impossible ici de défendre cette thèse de manière générale, et de passer en revue tous les domaines au sujet desquels il existe des points de rencontre féconds entre un type d'argumentation philosophique et des recherches en psychologie cognitive^[2]. Les plus discutés actuellement sont la théorie de l'esprit, la représentation et la planification de l'action, ou des hypothèses plus générales comme celle de la modularité. Je me limiterai ici à quelques exemples: le problème des rapports esprit-corps, la question de la conscience, et celle de la représentation mentale.

Commençons par le problème des rapports entre l'esprit et le corps, qui est à bien des égards le problème de la philosophie de l'esprit. La meilleure manière de le présenter de manière synthétique est, me semble-t-il, de remarquer que les trois propositions suivantes semblent plausibles, bien qu'elles se contredisent mutuellement :

- (1) les faits mentaux sont distincts des faits physiques
- (2) les faits mentaux ont des effets sur les faits physiques
- (3) et pourtant les faits physiques sont seuls suffisants pour causer des effets physiques

Un paralléliste niera (2) ; un dualiste interactionniste niera (3) ; un physicaliste ou un théoricien de l'identité esprit-cerveau niera (1). Il existe une position intermédiaire, explorée par les diverses formes de fonctionnalisme contemporaines, qui admettent que les propriétés mentales sont des propriétés fonctionnelles « réalisées » de manières diverses dans des structures physiques, ou par des auteurs comme Davidson (1980), qui entendent concilier les trois propositions en admettant que bien que les faits mentaux soient identiques à des faits physiques, nous ne pouvons pas espérer

complètement expliquer les premiers par les seconds. Ces auteurs défendent des formes non réductionnistes de physicalisme. Ils entendent néanmoins se ranger dans le camp physicaliste ou matérialiste dans la mesure où ils admettent la dépendance des faits mentaux par rapport aux faits physiques, au sens suivant : tout changement mental dépend, ou covarie, avec un changement physique, bien qu'il ne soit pas vrai que les mêmes différences mentales covarient avec ou dépendent des mêmes différences physiques. C'est cette relation que les philosophes appellent « survenance ». bien qu'il ne soit évidemment pas exhaustif, la plupart des solutions possibles au problème esprit-corps se distribuent sur le tableau suivant (inspiré de Van Gelder 1998):

<i>relation</i>	<i>1 parallélisme</i>	<i>2 dualisme causal</i>	<i>3 physicalisme faible</i>	<i>4 physicalisme fort</i>
identité	-	-	-	+
réduction	-	-	-	+
réalisation	-	-	+	+
survenance	-	-	+	+
interaction causale	-	+	+	+

Une bonne partie de la littérature en philosophie de l'esprit depuis une trentaine d'années a consisté à explorer une version ou une autre de ces positions, et en particulier celles du physicalisme faible (3), et les diverses versions de la survenance du mental sur le physique. La question ainsi posée consiste à chercher à réconcilier l'idée que les événements mentaux sont identiques à des événements physiques avec l'idée que les propriétés mentales et physiques des événements sont irréductiblement distinctes. La survenance des propriétés mentales sur les propriétés physiques est supposée être la relation appropriée. Mais comment peut-elle être plus qu'une simple relation de covariance, au sens où l'indiscernabilité de deux propriétés physiques implique l'indiscernabilité de deux propriétés mentales sans que les secondes se réduisent aux premières ? L'idée de survenance est incontestablement importante, et c'est un outil méthodologique important pour discuter une question fondamentale de la philosophie des sciences biologiques en général : celle de la nature de la réduction, et des variétés plus ou moins fortes de réductionnisme. Elle a également des liens

avec les métaphysiques émergentistes de la fin du siècle dernier. Mais il est douteux qu'elle nous permette de répondre au problème esprit-corps.

La notion de survenance est destinée à représenter l'une de nos intuitions quant à ce problème: celle d'une dépendance asymétrique entre le mental et le physique en dépit de la non réduction. Elle est, à ce titre essentielle quant il s'agit d'évaluer la relation entre psychologie et neurophysiologie, par exemple. Mais tel que le problème est posé, comme un problème portant sur l'ontologie des propriétés mentales et des propriétés physiques et leurs conditions abstraites de covariation, il est de nature exclusivement logique. Il est important de savoir si telle variété de survenance plus forte implique une autre, plus faible, mais ces problèmes ont plus à voir avec la manière dont on définit des notions comme celles de propriété, d'événement, d'instanciation d'une propriété, etc. Les relations en question sont purement formelles, comme le montre le fait que la relation de survenance peut aussi bien concerner la relation entre des propriétés morales et des propriétés naturelles. En ce sens la notion de survenance n'a pas un statut explicatif, mais seulement ontologique (Engel 1997). La question de la survenance du mental sur le physique ne peut perdre ce caractère formel que si l'on considère, en psychologie cognitive, des situations effectives de covariation des propriétés mentales par rapport à des propriétés physiques. On en a, me semble-t-il, un assez bon exemple avec la méthode de soustraction, utilisée notamment en imagerie cérébrale (Dehane 1997). Si on suppose qu'une certaine fonction cognitive F « réalise » la même propriété mentale (par exemple une tâche de mémoire) et si on veut voir comment cette propriété covarie (« survient sur ») avec, par exemple, le débit sanguin ou l'activité électromagnétique du cerveau, on forge des conditions expérimentales qui sont identiques en tout sauf pour l'étape de traitement mémoriel qu'on veut étudier, et on soustrait les mesures physiologiques effectuées dans les deux tâches. En ce sens on peut dire que la propriété mentale survient sur ces propriétés physiques. C'est déjà une application plus locale du concept que sous la forme précédente. Je voudrais suggérer que ce que les philosophes appellent *le* problème esprit-corps a toutes les

chances d'être insoluble s'il porte sur des relations métaphysico-formelles aussi générales que celles d'identité, de survenance, de réalisation et de causalité. Si l'on suppose que l'on a affaire, sans distinction, à des notions aussi homogènes que celles de *propriété mentale* ou *propriété physique*, le problème en question a peu de chances d'être résolu.

Et de fait, certains philosophes qui posent la question à ce degré de généralité soutiennent que le problème est insoluble. McGinn (1991) par exemple, identifie le problème esprit-corps au problème de la conscience. Il soutient de toute solution adéquate au problème consistera à identifier une propriété cérébrale P, et qu'en un sens nous disposons déjà d'une conception claire de la propriété en question. Il soutient cependant que même si nous pouvions avoir une analyse complète de la nature de P nous ne comprendrions toujours pas sa connexion avec les propriétés conscientes de notre expérience. Ni l'introspection, ni la connaissance physique du cerveau ne peuvent nous permettre de saisir le lien entre P et ce que nous appelons la conscience. Si l'on formule ainsi le problème, je suis tenté de donner raison à McGinn, dont la manière de voir ressemble assez à ce que l'on appelle parfois une théorie « du double aspect » : les phénomènes mentaux *sont* des phénomènes physiques, mais l'aspect sous lequel nous les saisissons nous interdit à jamais de le concevoir *comme* physiques. Ainsi posé, la question est bien insoluble. Mais c'est aussi parce que nous posons la question à ce degré de généralité que nous ne parvenons pas à comprendre quelle peut être la relation entre l'expérience consciente et les propriétés cérébrales. Le dilemme avancé par McGinn est insoluble si le seul choix qui nous est offert est entre une connaissance introspective de la conscience et la connaissance d'une propriété physique. La neuropsychologie, la psychopathologie, permettent d'envisager une troisième option, en étudiant, à nouveau par une méthode contrôlée de variations, les relations entre diverses lésions et des désordres de la conscience. Elles ne révèlent pas *la* propriété qui est responsable de l'expérience consciente, mais une trame de relations fonctionnelles propres à la cognition normale (et c'est pourquoi l'hypothèse de Crick, même si elle était correcte, serait

insuffisante). Là encore, parce qu'il n'est pas évident que nous ayons une propriété mentale homogène (la conscience), il n'est pas clair que la raison pour laquelle cette propriété nous semble indéchiffrable physiquement n'est pas liée à cette absence de spécificité.

Ce qui précède semble suggérer que plus une assertion portant sur des propriétés de l'esprit est générale et porte sur des entités homogènes, plus elle a de chances de rester *a priori*, c'est-à-dire, pour être charitable, au mieux vraie bien qu'invérifiable. Ce n'est pourtant pas ce que je veux suggérer. Je ne veux pas dire que les problèmes fondamentaux de la philosophie de l'esprit, comme celui de la relation corps-esprit, ou celui de la nature de l'intentionnalité sont des problèmes par définition insolubles, mettant en jeu des propositions invérifiables. Je pense au contraire que les théories forgées par les philosophes peuvent être testées, et évaluées, en particulier du point de vue de leur cohérence et de leur plausibilité empirique. Mais elles ne peuvent pas l'être directement. C'est pourquoi il nous faut essayer de déterminer un niveau intermédiaire de résolution des problèmes et d'interaction entre le niveau de description « philosophique » et le niveau de description psychologique (ou neuropsychologique, ou physiologique, selon les cas). Je voudrais essayer de montrer comment c'est possible au sujet de certains travaux sur la conscience.

Dans l'*Analytique transcendantale*, Kant soutient que l'unité de la conscience, pour un sujet, au sens de la possibilité que différents contenus de conscience soient associés à un seul centre, requiert que le sujet soit capable de faire l'expérience d'un monde objectif. C'est un type d'argument qu'on appelle couramment « transcendantal » : il part d'un fait d'expérience (la conscience est unifiée) et remonte à sa condition de possibilité (qu'il existe un monde objectif pour un sujet). Un autre argument qui conduit à la même conclusion peut être tiré de la célèbre objection de Lichtenberg au *cogito* cartésien : de quel droit pourrait-on dire « Je pense » en faisant référence à un sujet de nos pensées plutôt que « ça pense » qui ne contient aucune référence à un tel sujet? Mais l'objection de Lichtenberg ne marche pas. Car si l'on constate : « ça pense que P » , puis : « ça pense que Q », il devrait

pouvoir s'ensuivre : « ça pense que P et Q » ; mais cette pensée conjonctive ne peut pas surgir, car les deux pensées simples sont séparées, et peuvent ainsi surgir dans deux consciences distinctes. Mais elles doivent aussi pouvoir surgir dans la même conscience, si des pensées conjonctives de ce type sont simplement possibles. En d'autres termes, il doit pouvoir y avoir, si une unité de la conscience est possible, quelque chose comme un principe de transitivité : *s'il est pensé que P en i et s'il est pensé que Q en i, alors il est pensé que P et Q en i* (où 'i' est un indice qui dénote tout lieu, quel qu'il soit, où se produit la pensée en question). Supposons à présent que P et Q soient contradictoires. Il n'y aurait pas de problème à ce que P et Q soient pensés s'ils sont pensés dans deux consciences, ou à deux indices distincts, mais dans ce cas, il n'y aurait aucun moyen de déterminer si l'on a affaire à un seul ou à deux objets, c'est-à-dire si l'on a attribué à un même objet une propriété et sa contradictoire, ou deux propriétés contradictoires à deux objets différents. En d'autres termes, pour qu'il puisse y avoir conscience de deux objets distincts, il faut que le principe de transitivité ci-dessus soit possible, autrement dit qu'il puisse y avoir des épisodes de co-conscience, ou que la conscience ait une unité.^[3] Le lecteur psychologue qui aura suivi ce raisonnement jusqu'ici y aura reconnu le type même de raisonnement parfaitement *a priori* sur la conscience qui ne devraient avoir aucune incidence pour la recherche en psychologie expérimentale. Mais il semble qu'il y ait des cas où le principe de transitivité ci-dessus est violé, et donc où l'unité de la conscience est problématique. Un certain nombre de travaux sur les commissurotomies, comme ceux de Sergent (1990) semblent exemplifier cette possibilité. On projette des chiffres (par exemple '6' et '7') vers chaque hémisphère de patients ayant subi une commissurotomie, et on leur demande de les comparer. Dans une première tâche les sujets doivent répondre si deux chiffres sont les mêmes ou différents, et ne répondent pas mieux qu'au hasard. Dans une seconde tâche on leur demande si l'un des chiffres est plus élevé (ou plus grand) que l'autre. Mais ici ils trouvent la tâche facile et répondent de manière correcte. Et dans une troisième expérience, les patients sont

capables d'indiquer si les deux chiffres sont égaux ou si l'un est plus élevé que l'autre. Comment les patients peuvent-ils être incapables de répondre à des questions concernant l'identité de deux chiffres, tout en étant capables de répondre à des questions au sujet de leur égalité ou de leurs hauteurs ? Il est naturel de supposer qu'une information concernant la quantité comparée est partagée entre les deux hémisphères, mais pas concernant l'identité. Ces cas semblent illustrer directement des situations de non transitivité, ou de non unité de la conscience. La philosophe Susan Hurley (1996) qui examine ces cas, soutient qu'ils infirment la thèse d'une unité forte, mais pas celle d'une unité *faible*, de la conscience, qui expliquerait pourquoi les sujets, sans être capables de déterminer l'identité des chiffres qu'on leur présente, sont aussi capables de déterminer des propriétés partiellement constitutives de leur identité. Il s'ensuit qu'il est toujours vrai, comme le veut l'argument kantien, que cette unité faible de la conscience requiert la perception de propriétés objectives des objets.

Ce type d'argumentation est intéressant parce qu'il part d'une caractérisation purement *a priori* de l'unité de la conscience, pour non pas chercher directement à l'infirmer ou à la confirmer par des données psychologiques expérimentales, mais pour la modifier. Il faut cependant tempérer cette appréciation positive. La propriété de la conscience dont il a été question jusqu'ici est celle d'une expérience directe, rapportée dans des jugements. Mais il n'est pas évident que des cas comme ceux de Sergent portent réellement sur ce type de conscience : certaines réponses ne sont pas basées sur des états conscients, dans certains cas les patients devinent, et les échecs de la transitivité doivent porter sur des états qui se produisent au même moment. Pour qu'on puisse parler d'une unité de la conscience, forte ou faible, il faut que des conditions expérimentales bien spécifiques soient présentes.^[4] Ce n'est donc pas sans toutes sortes de précaution qu'on doit effectuer ces comparaisons entre des propositions philosophiques *a priori* et des constats expérimentaux. Quoiqu'il en soit, le raisonnement qui précède me paraît illustrer un type de démarche fécond et propre

à produire des interactions entre philosophie et psychologie. Il consiste à partir d'arguments transcendants au sujet d'une notion. Un argument transcendantal consiste à proposer la condition nécessaire d'un certain phénomène mental (par exemple ici la transitivité des états co-conscients). Cette condition en retour impliquera certaines thèses contrefactuelles sur ce qui devrait se passer si cette condition était absente (ici que la conscience devrait se trouver divisée). Habituellement les philosophes s'arrêtent là, et ces thèses contrefactuelles ne sont jamais vérifiées, sinon par des « expériences de pensée » qui reposent, comme leur nom l'indique, sur nos seules intuitions quant à ce qui est possible, ou concevable, dans telle situation. C'est ainsi qu'on raisonne, par exemple, au sujet de la notion d'identité personnelle, en envisageant diverses sortes de situations imaginaires dans lesquelles cette identité serait rompue (cf. par exemple Parfit 1984). Mais il y a moyen, en particulier quand on considère la neuropsychologie et la psychopathologie, d'envisager des structures de dissociation, et des contrastes entre cognition normale et cognition anormale, qui sont autant de vérifications possibles des thèses contrefactuelles en question. Dans certains cas, les arguments transcendantsaux seront vérifiés, dans d'autres ils seront révisés. Il est donc faux que ces arguments soient voués à rester purement *a priori*.

Je voudrais illustrer cette démarche sur un autre problème, extrêmement discuté en philosophie de l'esprit, celui de la nature de l'intentionnalité, de la représentation, ou de ce que les philosophes appellent quelquefois le « contenu » des états mentaux, ou plutôt un aspect de ce problème. L'une des hypothèses les plus séduisantes de la recherche en sciences cognitives est celle d'une structuration des représentations sous la forme d'un langage ou d'un code interne, ce que Fodor a appelé le « langage de la pensée ». Il est assez caractéristique de toute une littérature que cette hypothèse soit discutée la plupart du temps en termes purement *a priori*, c'est—à dire par des arguments supposés nous montrer pourquoi *il doit* y avoir un langage de la pensée, du type suivant (cf. par exemple Fodor 1988):

- (a) un trait propre aux pensées est leur *systematicité* (si on peut penser que Jean aime Marie, on doit pouvoir penser que Marie aime Jean), *leur productivité* (si on peut penser que Jean aime Marie, on doit aussi pouvoir penser que Raoul aime Marie, ou qu'il y a des gens qui s'aiment, etc.), et leurs *pouvoirs causaux* (si je pense que s'il neige, il fait froid, et si je viens à penser qu'il fait neige, je penserai en général qu'il fait froid)
- (b) ces trois propriétés s'expliquent parfaitement si l'on suppose que la pensée est encodée dans un langage, et la troisième s'explique parfaitement si l'on suppose que les structures physiques des symboles du langage mental ont un pouvoir causal sur d'autres structures physiques.

Mais il existe une autre option possible, parce que toutes les représentations ne sont pas nécessairement de type linguistique. Des cartes, des images, des diagrammes, peuvent être de représentations. Ils donnent, certes des informations en donnant *plus* d'information qu'une représentation linguistique, et sans avoir des articulations évaluables sémantiquement de manière discrète (chaque partie d'une carte contribue au contenu représentationnel de l'ensemble). Mais les cartes, les images, etc. ont aussi les propriétés de systématicité, de productivité et de causalité : on peut réarranger une carte pour qu'elle représente, par exemple, Paris au Sud de Lyon plutôt qu'inversement, un cartographe peut représenter autant de situations nouvelles que quelqu'un qui use de phrases, et les cartes sont, tout autant que les symboles d'un langage des structures physiques qui ont des pouvoirs causaux. L'idée qu'un grand nombre de représentations mentales pourraient être plus proches de cartes que de phrases d'un langage n'est donc pas moins justifiée *a priori* que l'hypothèse du langage mental. Elle s'accorde aussi bien avec des propriétés des états intentionnels sur lesquels insistent de nombreux philosophes, comme le *holisme* (la manière dont je crois que la neige est blanche ne correspond pas nécessairement à *une* croyance ayant ce contenu, mais à de nombreuses autres qui lui sont liées) qu'avec nombre de recherches familières en psychologie cognitive (les « cartes cognitives » chez les animaux et les humains, les « modèles mentaux » de Johnson-Laird en psychologie

du raisonnement, etc.). Mais le fait qu'un grand nombre de travaux en sciences cognitives semblent s'appuyer sur cette hypothèse n'est pas encore un exemple d'argument transcendantal. Pour qu'il y ait un argument de ce type, il faut que l'on dérive une condition nécessaire de toute représentation spatiale pour la vérifier ensuite sur des données expérimentales. Tout le monde connaît la propriété transcendantale de l'espace telle que l'expose Kant dans *l'Esthétique transcendantale* : toute représentation d'un univers objectif est nécessairement spatiale. John O'Keefe (1994), dans un essai qui s'adresse directement à ce problème kantien, soutient que les propriétés des cartes cognitives chez les rats qu'il a mis en évidence avec Nadel (O'Keefe and Nadel 1978) sont telles que ces représentations cérébrales ont une correspondance avec l'univers physique, que ces représentations spatiales sont antérieures à la notion d'un objet physique et indépendantes de lui, et que les cartes cognitives sont objectives, au sens où elles engendrent une représentation de l'environnement indépendante de l'organisme et de ses actions.

On répondra que ces discussions semblent amener les psychologues à se prononcer sur des notions ontologiques qui n'ont en fait pas grand chose à voir avec leur travail effectif. Mais ce type d'analyses n'est évidemment pas neuf en psychologie, en particulier en psychologie développementale : l'analyse de la causalité par Michotte, ou celle des nombres chez Piaget ne font pas autre chose, et des travaux de psychologie cognitive comme ceux de Spelke sur la perception des objets chez les enfants s'adressent directement à certains arguments transcendantsaux comme ceux de Quine (1960), qui soutenait que des catégories comme celles d'objet ou de substance émergent à partir d'une conceptualisation des objets en termes de masse et non d'individus discrets, et visent à les réfuter (Spelke 1990, Soja, Carey & Spelke 1991).

J'espère avoir indiqué une trame possible d'argumentation située au croisement de la philosophie et de la psychologie. La démarche est la même, à un point près, que celle de Kant : elle consiste à avancer certaines propriétés nécessaires

ou certaines conditions de possibilité de l'exercice de certaines facultés, de certains concepts ou de certaines représentations mentales, et à voir en quoi des travaux de psychologie cognitive, de neuropsychologie, ou d'autres disciplines permettent de les confirmer ou de les infirmer. Il s'agit d'une psychologie transcendantale, domaine intermédiaire, entre la psychologie pure ou rationnelle (dont l'exemple serait la discussion formelle du problème corps-esprit ou des relations de survenance) et la psychologie empirique. Cette psychologie détermine les conditions générales auxquelles doivent satisfaire des représentations, des concepts, ou des expériences pour qu'un esprit puisse accomplir des tâches cognitives. Mais cette psychologie transcendantale n'est-elle pas un sujet imaginaire ? Les authentiques kantien le penseront, car ils rejettent le « psychologisme », et identifient précisément le type d'enquête dont je viens de parler à la psychologie des facultés illustrée par la première édition de la *Critique de la raison pure* (c'est le « à un point près » mentionné ci-dessus). A cela il faut répondre : non, cette psychologie transcendantale n'est pas un sujet imaginaire, elle correspond précisément à un ensemble de travaux contemporains qui « naturalisent » le transcendantal kantien, comme j'ai essayé de le suggérer ici sur quelques exemples . J'anticipe l'objection des psychologues : peut-être ce sujet est-il intéressant pour les philosophes qui ont toujours quelques comptes à régler avec Kant) mais que peut-il nous apporter ? Ce que j'ai décrit n'est-il pas une simple variante de la méthode hypothético-déductive en sciences expérimentales : on propose une hypothèse, on en déduit des observations possibles, et on les teste ? Non, car les hypothèses « transcendantales » ne sont précisément pas des hypothèses empiriques, mais des hypothèses qui reposent sur la seule imagination conceptuelle. Mais cette imagination conceptuelle est ensuite confrontée aux données expérimentales. Je ne soutiendrai pas que c'est le seul type d'interaction fécond entre philosophie et psychologie. Il y a bien sûr une manière plus traditionnelle de travailler, qui consiste, pour le philosophe, à s'immerger dans les théories de la psychologie contemporaine, et à essayer de dégager de ces travaux les théories de

l'esprit les plus plausibles. Je n'ai pas voulu exclure cette manière de travailler. Mais j'espère avoir suggéré ici qu'il y en a une autre.

Pascal Engel

professeur de philosophie, Université de Paris-IV Sorbonne, membre du GEMAS (CNRS), membre associé du CREA. Thèmes de recherche : philosophie de la logique et du langage, philosophie de l'esprit et de la psychologie.

NOTES

REFERENCES

Bermudez, J. 1995 a « Transcendental arguments and psychology », *Metaphilosophy*, 26, 4, 379-401

Davidson, D. 1980 *Essays on Actions and Events*, Oxford, Oxford University Press, tr.fr. *Actions et événements*, Paris, PUF 1993

Dehane, S. 1997 « La méthode de soustraction », in Dehane, S. ed, *Le cerveau en action*,

imagerie cérébrale fonctionnelle et psychologie cognitive, Paris, PUF, 1997

Engel, P. 1996 *Philosophie et psychologie*, Paris, Gallimard, Folio

1997 « Peut-on résoudre le problème de la causalité mentale? » in J.C. Bourdin

ed. *Les matérialismes*, ed. Kimé, Paris, 197-218

Fodor, J. 1988, *Psychosemantics*, Cambridge Mass, MIT Press

Goldman, A. ed. 1993 *Readings in Philosophy and Cognitive Science*; Cambridge, Mass.

MIT Press

Hurley, S. 1994 « Unity and Objectivity », in Peacocke 1994, 49-78

Kim, J. 1993 *Supervenience and Mind*, Cambridge, Cambridge University Press

Marcel, A. 1994 « What is relevant to the Unity of Consciousness ? », in Peacocke 1994,

- McGinn, C. 1991 *The problem of Consciousness*, Oxford, Oxford University Press
- O'Keefe, J. and Nadel, L 1978 *The Hippocampus as a Cognitive Map*, Oxford, Oxford University Press
- O'Keefe, J. (1994) « « Cognitive Maps, Time and Causality », in Peacocke 1994, 35-48
- Parfit, D. 1984, *Reasons and Persons*, Oxford, Oxford University Press
- Peacocke, C. 1994, ed. *Objectivity, Simulation and the Unity of Consciousness*, Oxford, Oxford University Press
- Sergent, J. (1990) « furtive Incursions into Bicameral Minds », *Brain*, 113, 537-568
- Spelke, E. 1991 « Origins of visual knowledge », in D. Osherson, ed. *An invitation to cognitive science, vol. 2, Cambridge, Mass*, repr. in Goldman 1993, 448-460
- Spelke, Soja & Carey, 1991 « Ontological Categories Guide Young children's Instructions of Word Meaning », *Cognition*, 38, 179-211, repr. in Goldman 1993, 461-480
- Van Gelder 1987 « Monism, Dualism, Pluralism », *Mind and Language*, 13, 156-75
- Williams, B. 1978 *Descartes*, Penguin Books, Hammondsworth, London

[1] L'usage de ces diverses terminologies n'est en général pas neutre.. En France, on entend souvent ceux qui éprouvent vis à vis du mot « esprit » une certaine répulsion (voire une répulsion certaine) dire que le terme de « philosophie cognitive » est meilleure. « Philosophie de la psychologie » est rarement utilisé, pas plus que « psychologie philosophique » qu'on trouve souvent en anglais. A mes yeux ces différences de terminologie ont peut d'importance, et la seule raison que je vois de refuser le terme « philosophie de l'esprit » est qu'on souscrit à une version de l'éliminativisme.

[2] J'ai déjà proposé cette conception dans Engel 1996, en particulier ch.1, p.111-132, et ch.IV passim

[3] Tout le raisonnement qui précède est celui de B. Williams (1978, p.97), tel qu'il est repris par Hurley (1996).

[4] C'est ce qu'objecte Marcel 1996 à Hurley (1996)