

## Philosophiques

philosophiques

Hilary Putnam, *Représentation et réalité*, traduction française par Claudine Engel-Tiercelin, Paris, Éditions Gallimard, 1990, 226 pages.

Paul Bernier

Volume 18, numéro 2, automne 1991

URI : <https://id.erudit.org/iderudit/027165ar>

DOI : <https://doi.org/10.7202/027165ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer ce compte rendu

Bernier, P. (1991). Compte rendu de [Hilary Putnam, *Représentation et réalité*, traduction française par Claudine Engel-Tiercelin, Paris, Éditions Gallimard, 1990, 226 pages.] *Philosophiques*, 18(2), 191–195.  
<https://doi.org/10.7202/027165ar>

Hilary Putnam, *Représentation et réalité*, traduction française par Claudine Engel-Tiercelin, Paris, Éditions Gallimard, 1990, 226 pages.  
par Paul Bernier

Le propos de Putnam est principalement de rejeter la théorie computationnelle de l'esprit selon laquelle « l'ordinateur est le modèle qui convient pour l'esprit » (p. 13). Dans le dernier chapitre, il s'explique sur la position métaphysique qui l'amène à rejeter cette théorie qu'il fut l'un des premiers à défendre. Ainsi, c'est parce qu'il a abandonné le réalisme scientifique qu'il défendait jadis que Putnam cherche à montrer l'impossibilité (en principe) de réduire la notion d'intentionnalité dans le cadre d'une théorie computationnelle. De plus, l'impossibilité d'une telle réduction ne serait pas une raison d'adopter une thèse éliminationniste face à l'intentionnalité. Il consacre

un chapitre (ch. 4) à tenter de montrer que même si la notion d'intentionnalité repose en dernière analyse sur nos intuitions de la psychologie du sens commun, elle ne devrait pas pour autant être éliminée, pas plus que ne doit l'être la notion de vérité en logique, qui repose également sur nos intuitions du sens commun. Ce chapitre est assez indépendant du reste du propos de Putnam qui vise principalement à argumenter contre la possibilité d'une réduction de la notion d'intentionnalité dans un cadre computationnel ou fonctionnaliste.

Les principaux arguments s'appuient essentiellement sur trois principes, soit (1) le holisme de la signification, (2) le fait que la signification est une notion partiellement normative et (3) la contribution de l'environnement physique et socio-linguistique à la signification. Putnam a largement discuté le troisième principe dans d'autres écrits et il reprend ici essentiellement les mêmes idées. Par ailleurs, on doit déplorer le fait qu'il ne développe pas de façon très détaillée les deux premiers principes. Pour ce qui concerne la thèse du holisme sémantique, Putnam s'appuie sur la thèse quinième du holisme épistémologique. On doit noter cependant qu'il n'est pas évident que cette thèse épistémologique puisse être directement transposée au niveau de l'analyse sémantique, surtout lorsqu'il s'agit d'analyser les notions de croyance et de désir qui sont le fait de la psychologie du sens commun.

Putnam développe deux arguments principaux. Le premier (ch. 2) consiste à montrer que la notion d'intentionnalité ne peut être réduite aux représentations mentales internes postulées par la théorie computationnelle de l'esprit. Ce serait en raison de la division du travail linguistique et de la contribution de l'environnement physique (troisième principe) que la signification des termes ne peut être déterminée par les représentations mentales qui y sont associées. La signification de certains termes est sous-déterminée par les représentations mentales qui en sous-tendent l'utilisation, puisque deux termes qui ont des significations distinctes (par exemple « hêtre » et « orme » ou encore « eau » utilisé sur la Terre et sur Terre-jumelle) peuvent néanmoins correspondre à une seule représentation mentale; schématiquement: même représentation mentale  $\Rightarrow$  significations différentes. Putnam consacre un chapitre (ch. 3) à rejeter certaines objections qui ont été soulevées contre cet argument qui repose sur la thèse du caractère essentiellement externe du contenu des représentations mentales. Il rejette les objections de Jerry Fodor et de Ned Block, qui consistent à distinguer deux notions de contenu, soit le contenu large et le contenu étroit. Selon ces objections, les facteurs externes ne détermineraient que le contenu large des représentations mentales; aussi celles-ci ne seraient pas pour autant dépourvues d'intentionnalité, puisqu'elles auraient également un contenu étroit. Quoi qu'il en soit de la force des contre-arguments de Putnam, on doit noter qu'il ne discute pas la notion de contenu étroit lorsqu'elle est conçue comme étant une fonction du contexte au contenu large. Putnam

est certes au fait de cette théorie du contenu étroit que Fodor a développée dans *Psychosemantics* — puisqu'il en fait mention —, mais il s'attaque plutôt à une théorie du contenu étroit que Fodor n'a jamais soutenue assez sérieusement pour en faire la publication. On doit également noter que la notion de contenu étroit est l'objet de nombreux travaux, et qu'elle prend par conséquent des formes variées chez les philosophes qui ont développé cette notion. Aussi, même si Putnam donne l'impression de rejeter les principales objections à son argument externaliste, il ne faudrait pas s'y méprendre.

Aux chapitres cinq et six, Putnam développe une autre argumentation qui a une portée plus générale, en ce sens qu'elle s'appliquerait même à ce qu'on pourrait appeler un fonctionnalisme large, c'est-à-dire à un fonctionnalisme qui individuerait les états computationnels en tenant compte des facteurs externes. Putnam souligne d'abord que toute théorie qui prétend produire une réduction de la notion de référence devrait être en mesure (a) de définir une relation *R* qui est la notion réductrice, (b) cette notion devrait obéir aux « lois » auxquelles la notion de référence obéit dans le langage ordinaire, et (c) elle devrait expliquer les mêmes effets que la notion de référence arrive à expliquer. Or le défi que Putnam pose à la théorie computationnelle est de produire ou bien une définition explicite ou bien une définition récursive d'une telle relation *R* qui satisfasse ces trois desiderata. Putnam tente de montrer que cela est impossible même en principe. Ainsi, contrairement à l'argument précédent, celui-ci vise à montrer non pas la sous-détermination des significations par les représentations mentales, mais plutôt l'indétermination des états physico-computationnels qui peuvent sous-tendre des termes synonymes. En d'autres termes, Putnam soutient que deux expressions qui, dans nos pratiques interprétatives, sont généralement acceptées comme synonymes pourraient néanmoins être sous-tendues par un nombre infini d'états physico-computationnels distincts. Ce que schématiquement on peut exprimer ainsi: même signification états  $\Rightarrow$  physico-computationnels différents (et en nombre infini). Cet argument s'appuie sur les deux premiers principes introduits au premier chapitre. Le holisme de la signification implique que deux personnes qui utilisent deux expressions jugées synonymes dans leurs pratiques interprétatives ont cependant des croyances distinctes relativement à cette expression. Puisque chaque individu a acquis la maîtrise des expressions linguistiques qu'il utilise dans des situations particulières distinctes, deux expressions jugées synonymes sont néanmoins associées chez deux individus à des croyances distinctes. C'est-à-dire que deux individus qui utilisent une expression qu'ils jugent synonymes ont cependant certaines croyances divergentes relativement à cette expression. Si pour interpréter deux expressions comme synonymes, il fallait qu'elles soient sous-tendues par des réseaux de croyances identiques, alors deux expressions ne pourraient jamais être interprétées comme synonymes. Ainsi pour que Putnam puisse soutenir son argument de la réalisabilité physico-computationnelle potentiellement infinie (des états qui sous-tendent les expressions jugées synonymes),

il doit également soutenir qu'en principe deux expressions linguistiques synonymes peuvent être sous-tendues par des réseaux doxastiques différents (en nombre potentiellement infini). Le fait que deux individus ne partagent jamais exactement les mêmes réseaux doxastiques n'empêche aucunement que, dans leurs pratiques interprétatives, ils tiennent pour synonymes certains termes qu'ils emploient respectivement. Et c'est justement à cet égard que le second principe (normativité) intervient. L'argument vise donc à montrer que même si une théorie computationnelle de l'esprit est restreinte par un principe de normativité — comme l'est une théorie de l'interprétation lorsqu'il s'agit de rendre compte de la synonymie —, cela ne la restreint pas suffisamment, de sorte que cette contrainte laisse ouverte la possibilité qu'un nombre infini de systèmes physico-computationnels (et par conséquent d'états computationnels) sous-tendent des termes synonymes.

Au chapitre six Putnam développe cette argumentation en tentant de montrer que même si la théorie computationnelle était limitée par une « théorie psychologique idéale » ou par la psychologie du sens commun — interprétée comme une théorie (comme c'est le cas chez David Lewis) —, cela ne suffirait pas à restreindre l'espace (infini) des réalisations possibles d'états computationnels sous-jacents aux expressions synonymes. Pour qu'il y ait réduction de la théorie psychologique (idéale ou du sens commun), il doit y avoir un isomorphisme entre d'une part des classes de systèmes physico-computationnels et d'autre part les états mentaux de la théorie psychologique. Et encore une fois, Putnam soutient que la classe des systèmes physico-computationnels qui réalisent un des états mentaux de la théorie psychologique est ouverte (*open-ended*).

Bien que cet ouvrage de Putnam ne convaincra sûrement pas les philosophes qui sont sympathiques à la théorie computationnelle de l'esprit — puisque ses arguments semblent vulnérables face à certaines objections qu'il ne considère pas —, ce livre a au moins l'intérêt de brosser un tableau général des problèmes que rencontrent les théories de l'esprit qui s'inspirent des modèles computationnels utilisés en sciences cognitives. Concernant le second argument de Putnam (ch. 5 et 6), son propos met bien en évidence le problème qui consiste à imposer certaines contraintes pour limiter la classe des systèmes physico-computationnels qui réalisent les états mentaux. Il est curieux cependant que Putnam se soit limité à ne considérer que les contraintes imposées ou bien par une théorie psychologique idéale (mal définie, comme il l'admet lui-même), ou bien par la théorie psychologique du sens commun (guère mieux définie), quand en fait il existe certaines théories empiriques qui imposent des contraintes empiriques sur les phénomènes psychologiques, comme la théorie de la vision de Marr, ou encore certains modèles de l'apprentissage du langage. Si le but d'une théorie de l'esprit est de rendre compte de la Rationalité, telle qu'elle se manifeste dans nos pratiques interprétatives, alors on peut concéder que les modèles computationnels font face

à des difficultés de principe insurmontables. Mais à notre avis, exiger d'une théorie computationnelle qu'elle constitue une réduction d'une théorie de nos pratiques interprétatives (comme Putnam le fait de façon implicite tout au long de son ouvrage), c'est mal comprendre les buts et les motivations des théories computationnelles. Quant au réalisme interne que Putnam reprend et développe au dernier chapitre, il semble s'agir d'une position de compromis très difficile à soutenir.

*Département de philosophie  
Université de Montréal*

