Théorie de l'esprit ou simulation

Anne Reboul

▶ **To cite this version:**

Anne Reboul. Théorie de l'esprit ou simulation : l'apport des études développementales. 2005. halshs-00003838

HAL Id: halshs-00003838
https://shs.hal.science/halshs-00003838

Preprint submitted on 7 Feb 2005

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.

(à paraître), in *Confrontations psychiatriques*

Théorie de l'esprit ou simulation: l'apport des études développementales[1]

Anne Reboul, Institut des Sciences Cognitives, CNRS-UMR 5015, 67 bd Pinel, 69675 Bron cedex, reboul@isc.cnrs.fr

Everything should be made as simple as possible, but not simpler.
Albert Einstein

Résumé : Sur la base des indications et des définitions données par B. Pachoud dans son intervention, je dresserai rapidement un panorama de l'acquisition parallèle du langage et de la théorie de l'esprit et j'indiquerai les conclusions que l'on peut tirer de ces données pour le débat entre la vision "théorique" et la vision "simulationniste" de la théorie de l'esprit. Spécifiquement, les débuts de la théorie de l'esprit (le développement de la reconnaissance de l'agentivité et de la capacité d'attention partagée) sont en place lorsque l'enfant commence à acquérir le lexique. Ces capacités ne sont clairement pas de nature théorique. La capacité à passer le test de la fausse croyance (acquise entre trois et quatre ans) semble dépendre de la maîtrise de concepts et de structures linguistiques. Elle paraît de nature plus théorique que simulationniste. La théorie de l'esprit repose donc probablement sur une combinaison de simulation et de connaissances plus théoriques. Cette vision des choses pourrait expliquer certaines particularités de l'acquisition du langage par les autistes et les patients atteints du syndrome d'Asperger.
Mots-clé : théorie de l'esprit, simulation, acquisition du langage, autisme, syndrome d'Asperger

1. Introduction

Comme l'a rappelé Bernard Pachoud, la théorie de l'esprit, la capacité à attribuer à autrui des états mentaux et à prédire ou à expliquer son comportement sur cette base, est expliquée dans la littérature à partir de deux notions : la notion de simulation et la notion de théorie. Selon la première, très grossièrement, la théorie de l'esprit reposerait sur la capacité à se simuler dans la situation d'autrui, puis à lui attribuer les résultats ainsi obtenus. Il y a non pas une mais deux versions de la seconde : certains (Fodor 1992, Scholl & Leslie 1994) considèrent que la théorie de l'esprit est une théorie innée, de nature plus ou moins modulaire ; d'autres (Perner 1991, Gopnik & Meltzoff 1997) considèrent que c'est une théorie construite au fur et à mesure du développement, et dont le développement est contraint par l'acquisition de nouveaux concepts (comme, par exemple, la *croyance* ou la *représentation*). Pour diverses raisons, je pense que la version "acquise" de la théorie de l'esprit comme théorie est plus intéressante et plus convaincante que sa version innée et je ne considérerai donc ici que la vision "simulationniste" de la théorie de l'esprit et la vision selon laquelle c'est une théorie acquise.

Sans prétendre le moins du monde trancher le débat, on peut néanmoins se demander s'il est bien posé lorsqu'il s'agit de décider si la théorie de l'esprit est une simulation ou une théorie. On peut très bien penser (et c'est la position que j'adopterai) que la théorie de l'esprit n'est pas un monolithe (d'où ma réticence à adopter la position innéiste sur la théorie de l'esprit comme théorie), mais une

[1] Cet article est dédié à la mémoire de mon frère, Dominique Reboul (1960-2002).

capacité composite qui met en jeu aussi bien la tendance à la simulation que des processus inférentiels complexes guidés par une théorie au moins en partie acquise. Cette position mixte a été notamment défendue par Josef Perner (1996) et me semble convaincante notamment lorsqu'elle est confrontée aux données développementales sur l'acquisition aussi bien du langage que de la théorie de l'esprit.

2. Les données développementales

Dans un ouvrage récent, le psychologue Paul Bloom (2000) défend la thèse selon laquelle l'acquisition du lexique[2] se fait sur la base de la théorie de l'esprit : c'est cette capacité qui permet à l'enfant d'identifier ce dont on lui parle en décelant les intentions de son interlocuteur. Cette position, bien qu'elle soit brillamment défendue par son auteur, se heurte néanmoins à plusieurs objections : le test de la fausse croyance (pierre de touche de la maîtrise de la théorie de l'esprit) n'est généralement passé avec succès qu'entre 3 et 4 ans alors que l'acquisition du lexique débute vers 12 mois ; ensuite, Glüer et Pagin (2003) rappellent que certains autistes, chez lesquels la théorie de l'esprit est généralement considérée comme dysfonctionnelle, apprennent néanmoins à parler ; qui plus est, l'acquisition du langage, si elle ne suit pas le calendrier habituel chez les autistes, semble normale chez les patients atteints du syndrome d'Asperger et chez qui la théorie de l'esprit serait également dysfonctionnelle. Par ailleurs, certains auteurs défendent, sur une base expérimentale, l'hypothèse d'un lien entre capacité à passer le test de la fausse croyance et maîtrise linguistique, bien que le débat reste vif sur le fait que la maîtrise linguistique en question porte sur des facteurs purement syntaxiques, purement sémantiques ou syntaxiques et sémantiques (cf. notamment Yun Chin & Bernard-Opitz 2000, de Villiers & Pyers 2002, Ruffman *et al.* 2003). Enfin, certaines expériences récentes (cf. Kamawar & Olson 2000, Robinson & Apperly 2001) semblent indiquer que la maîtrise des contextes opaques n'est acquise qu'un an plus tard, vers 5 ans.

Je voudrais très rapidement décrire les expériences citées plus haut de la fausse croyance et des contextes opaques. L'expérience de la fausse croyance consiste à présenter aux enfants une saynète jouée généralement par des poupées. La poupée Sally a une bille qu'elle met dans son panier. Puis elle sort et, en son absence, la poupée Ann prend la bille et la met dans sa boite. Lorsque la poupée Sally revient, on demande aux enfants où elle va (d'abord) aller chercher la bille. Les enfants commencent à donner la réponse correcte ("dans le panier") entre trois et quatre ans. L'expérience des contextes opaques s'explique mieux à partir d'exemples linguistiques:

(1) Aristote était le précepteur d'Alexandre le Grand.
(2) Aristote était l'élève de Platon.
(3) May croit qu'Aristote était le précepteur d'Alexandre le Grand
(4) May ne croit pas qu'Aristote était l'élève de Platon.
(5) L'élève de Platon (l'un d'entre eux) était le précepteur d'Alexandre le Grand.
(6) May croit que l'élève de Platon était le précepteur d'Alexandre le Grand.

Les phrases (1) à (4) sont les prémisses (vraies). La phrase (5) est une inférence que l'on peut légitimement tirer des phrases (1) et (2). La phrase (6), à l'inverse, n'est pas une inférence que l'on peut légitimement tirer de (3) et (4). Passer le test des

[2] L'acquisition des structures syntaxiques est probablement le fruit d'un processus différent, dont on considère souvent qu'il est en partie inné (cf. Chomsky 1981).

contextes opaques[3], suppose que l'on soit capable, à partir des phrases (3) et (4), de rejeter l'inférence en (6)[4]. On peut considérer que le passage du test des contextes opaques, par la maîtrise qu'il révèle du contexte de *croyance* indique la maîtrise de la théorie de l'esprit[5].

Ainsi, on le voit, loin que la théorie de l'esprit soit en place au départ de l'acquisition du lexique (vers 12 mois), sa maîtrise n'est réalisée que vers 5 ans. Pour autant, l'hypothèse de Bloom s'appuie sur des arguments intéressants et qui ne peuvent être ignorés. On peut donc supposer que la théorie de l'esprit est précédée par d'autres mécanismes qui pourraient assurer tout à la fois son acquisition et l'acquisition du lexique. Dans son ouvrage de 1995, Baron-Cohen décrit un certain nombre de mécanismes dont la mise en place est précoce. Il s'agit du détecteur d'intentionnalité (*ID*), du détecteur de direction oculaire (*EDD*) et, sur la base des deux mécanismes précédents, du mécanisme d'attention partagée (*SAM*). C'est, selon lui, sur cette base que se construirait la théorie de l'esprit (*ToM*).

Le tableau de l'acquisition conjointe de la théorie de l'esprit et du lexique, sur la base des données proposées par Baron-Cohen, serait donc le suivant :

Age	Acquisition du lexique	Acquisition de la T. de l'esprit
0-9 mois		ID et EDD
9-18 mois	40 mots	SAM
24 mois	311 mots	Développement de ToM
30 mois	574 mots	Développement de ToM
48 mois	Développement du lexique	Test de la fausse croyance
60 mois	Développement du lexique	Test des contextes opaques

Tableau 1 : Chronologie du développement conjoint du lexique et de la théorie de l'esprit

Ainsi, l'acquisition de la théorie de l'esprit est loin d'être complètement réalisée lors du déclenchement de l'acquisition lexicale. En revanche certains mécanismes qui conditionnent à la fois l'acquisition de la théorie de l'esprit et celle du lexique sont en place. On pourrait considérer que ces mécanismes précoces, bientôt suivis du mécanisme d'attention partagée, et qui sont clairement de nature non-théorique, sont une des bases sur lesquelles se construit la théorie de l'esprit comme une théorie, une autre base étant l'acquisition, par la voie linguistique de concepts comme la *représentation* et la *croyance*[6].

---

[3] Ainsi appelés parce que les verbes d'attitude propositionnelle (*croire, savoir* , etc.) contraignent les inférences possibles, comme on vient de le voir.

[4] Le test proposé aux enfants prend l'exemple d'un jouet qui est aussi un cadeau. Le principe est exactement similaire à celui décrit ci-dessus.

[5] Ceci ne signifie pas qu'elle ne continue pas à se développer par la suite. Simplement, les principes fondamentaux en sont acquis.

[6] On notera que cette position n'interdit pas de penser que l'acquisition de la théorie de l'esprit passe aussi par des processus socio-affectifs comme ceux décrits par Trevarthen (1977) ou Sigman *et al.* (1986). Comme l'a montré Damasio (1994, 1999, 2003), l'émotion et la cognition sont parfaitement compatibles et s'il est un domaine où l'on peut s'attendre à voir l'apprentissage guidé par l'émotion, c'est bien la théorie de l'esprit.

3. Simulationnistes ou … non-théoriques

Quelques objections simples viennent à l'esprit. Certes, ID et EDD ne sont pas des mécanismes théoriques. Mais il n'y a pas de raison de penser que ce soit des mécanismes simulationnistes. La double équation simple qui voudrait que théorique = non-simulationniste et que non-théorique = simulationniste est donc battue en brèche. Si ID et EDD ne sont pas des mécanismes théoriques, ce ne sont pas davantage des mécanismes simulationnistes : plutôt, ce pourraient être les bases tout à la fois de processus de simulation et de processus théoriques qui sous-tendent conjointement la théorie de l'esprit. En d'autres termes, le fait que ID et EDD sous-tendent l'acquisition de la théorie de l'esprit comme théorie n'est en rien une preuve du fait que la théorie de l'esprit soit le résultat tout à la fois de processus simulationnistes et de processus théoriques. Je voudrais très rapidement indiquer les développements récents des travaux de Perner (cf. Dienes & Perner 1999, Perner 2003 et Perner & Dienes 2003). Dans ces travaux, les auteurs tirent une distinction entre les processus implicites quant à la prédication (*predication-implicit*) et les processus explicites quant à la prédication (*predication-explicit*). Les premiers ont pour exemples prototypiques l'initiation et le contrôle autonome (indépendant de toute représentation explicite) et automatique des actions ; les seconds ont pour exemples paradigmatiques l'acquisition de nouveaux schémas d'action via des instructions verbales, la planification, le raisonnement ou le test d'hypothèse.

Perner (1996) indique explicitement qu'un processus de simulation est un processus implicite quant à la prédication, alors qu'un processus inférentiel déclenché par la théorie est un processus explicite quant à la prédication. Plus précisément, un processus simulationniste consiste à adopter le point de vue (les connaissances ou les croyances d'autrui) et à en tirer des conclusions que l'on attribue à autrui sur la base de mécanismes cognitifs indépendants. Dans un tel processus, les prémisses sur lesquels travaillent les processus cognitifs ne sont pas représentées comme croyances et les inférences basées sur la théorie de l'esprit (comme théorie) ne sont pas autorisées. Un exemple d'un processus simulationniste est celui d'un individu A auquel on demande si un autre individu B, parlant la même langue maternelle, considérerait une phrase donnée de cette langue comme grammaticale ou agrammaticale. Pour répondre à cette question, A ne va pas faire appel à une théorie sur la façon dont B (ou autrui en général) détermine le caractère grammatical ou non d'une phrase. Il va se contenter de considérer la phrase en question, de se demander s'il la trouve ou non grammaticale et d'attribuer sa propre réponse à B. On remarquera qu'un faible élément de théorie entre dans la détermination de ce sur quoi travaille le processus : c'est le degré plus ou moins grand de similitude entre le sujet et autrui.

De façon générale, les tenants de la mixité plutôt que de la séparation tendent à considérer que les processus de simulation pourraient jouer un rôle dans la détermination du contenu, mais pas dans celui des inférences basées sur les relations entre attitudes propositionnelles (croyance, connaissance, désir, etc.). On pourrait donc ainsi considérer que la capacité de simulation (implicite quant à la prédication, i.e. aveugle quant aux attitudes propositionnelles) est acquise très tôt sur la base de ID, EDD et SAM si tant est qu'elle ne soit pas au moins partiellement innée, tandis que la théorie (de la théorie de l'esprit) fait l'objet d'un apprentissage long, qui s'appuie sur des processus explicites quant à la prédication et qui dépend de ce fait, en grande partie, de l'apprentissage du langage.

4. En conclusion : l'autisme, une faillite de la théorie ou de la simulation ?

Ainsi, la théorie de l'esprit mettrait en jeu la simulation et la théorie. Dès lors, et si l'on accepte la doxa selon laquelle l'autisme et les syndromes qui lui sont associés viennent au moins en partie d'un dysfonctionnement de la théorie de l'esprit, faut-il considérer que l'autisme est lié à un dysfonctionnement du volet simulationniste ou du volet théorique de la théorie de l'esprit ? A l'heure actuelle, il est extrêmement difficile de répondre à cette question, étant donné que distinguer expérimentalement ce qui ressortit à la théorie et ce qui ressortit à la simulation est loin d'être simple (cf. Perner *et al.* 1999, Perner & Kühberger ?, 2003). On peut néanmoins écarter une hypothèse. Le déficit en théorie de l'esprit des autistes ne provient pas d'un dysfonctionnement des mécanismes de base que sont ID et EDD (cf. Baron-Cohen 1995). Cependant, le même auteur détecte un problème au niveau de SAM, le mécanisme d'attention partagée. Le problème pourrait donc être transversal à la théorie de l'esprit et transgresser la frontière, en tout état de cause ténue, entre théorie et simulation.

Quelques mots pour finir : les autistes qui ont acquis le langage n'en conservent pas moins un certain nombre de problèmes de nature pragmatique (liés à l'usage du langage) ainsi que des difficultés en théorie de l'esprit. Si la théorie de l'esprit est (au moins pour sa composante théorique) liée à la maîtrise du langage, on peut se demander pourquoi leurs problèmes subsistent. La réponse à cette question repose sur plusieurs faits ou hypothèses : d'une part, leur acquisition du langage ne s'est pas faite selon le calendrier et par les processus habituels ; d'autre part, leur maîtrise du langage reste inférieure à celle de sujets d'âge mental équivalent ; ensuite, leurs problèmes pragmatiques et leurs difficultés en théorie de l'esprit peuvent provenir de difficultés sur la part simulationniste de leur théorie de l'esprit.

Quoi qu'il en soit, j'ai cherché à montrer que les processus qui sous-tendent la théorie de l'esprit sont proprement plus complexes qu'on ne le considère généralement et que les questions autour de la théorie de l'esprit ne peuvent se résumer à un simple choix entre simulation et théorie. De ce fait, la nature du trouble autistique est lui aussi probablement plus complexe, comme le sont l'acquisition même de la théorie de l'esprit et celle du langage.

Références

Baron-Cohen, S. 1995. *Mindblindness. An essay on autism and theory of mind*, Cambridge, MA, The MIT Press.

Bloom, P. 2000. *How Children Learn the Meanings of Words*, Cambridge, MA, The MIT Press.

Chomsky, N. 1981. *Lectures on Government and Binding*, Dordrecht, Foris.

Damasio, A. (1994) *L'erreur de Descartes. La raison des émotions*, Paris, Odile Jacob.

Damasio, A. (1999) *Le sentiment même de soi. Corps, émotions, conscience*, Paris, Odile Jacob.

Damasio, A. (2003) *Looking for Spinoza. Joy, Sorrow and the Feeling Brain*, Orlando, Florida, Harcourt.

De Villiers, J.G. and Pyers, J.E. (2002), Complements to cognition: a longitudinal study of the relationship between complex syntax and false-belief understanding, *Cognitive Development* 17, 1037-1060.

Dienes, Z & Perner, J. (1999) A theory of implicit and explicit knowledge, *Behavioral and Brain Sciences* 22, 735-808.

Fodor, J. 1992. Discussion: A theory of the child's theory of mind. *Cognition 44*, 283-296.

Glüer, K. & Pagin, P. (2003) Meaning theory and autistic speakers, *Mind and Language* 18/1, pp. 23-51.

Gopnik, A. & A.N. Meltzoff 1997. *Words, Thoughts and Theories*, Cambridge, MA, The MIT Press.

Kamavar, D. and Olson, D.R. (2000), Children's Representational Theory of Language: Problems of Opaque Contexts, *Cognitive Development* 14, pp. 531-548.

Perner, J. & Dienes, Z. (2003) Developmental aspects of consciousness: how much theory of mind do you need to be consciously aware?, *Consciousness and cognition* 12, 63-82.

Perner, J. & Kühberger, A. (à paraître) Framing and the theory-simulation controversy: predicting people's decision, *Mind and society.*

Perner, J. & Kühberger, A. (à paraître) Putting philosophy to work by making simulation theory testable: the case of endowment, in Kanzian, C., Quitterer, J. & Runggaldier, E. (eds), *Persons. An interdisciplinary approach. Personen. Ein interdisziplinärer Dialog*, Vienne, Obvahpt.

Perner, J. (1996) Simulation as explicitation of predication-implicit knowledge about the mind: arguments for a simulation-theory mix, in Carruthers, P. & Smith, P.K. (eds), *Theories of theories of mind*, Cambridge, Cambridge University Press, 90-104.

Perner, J. (2003) Dual control and the causal theory of action: the case of non-intentional action, in Eilan, N. & Roessler, J. (eds), *Agency and self-awareness*, Oxford, Oxford University Press, 218-243.

Perner, J. 1991. *Understanding the Representational Mind*, Cambridge, MA, The MIT Press.

Perner, J., Gschaider, A., Kühberger, A. & Schrofner, S. (1999) Predicting others through simulation or by theory? A method to decide, *Mind and Language* 14, 57-79.

Robinson, E.J. and Apperly, I.A. (2001) Children's difficulties with partial representations in ambiguous messages and referentially opaque contexts, *Cognitive Development* 16, pp. 595-615.

Ruffman, T., Slade, L., Rowlandson, K., Rumsey, C. and Garnham, A. (2003) How language relates to belief, desire and emotion understanding, *Cognitive Development* 18, pp. 139-158.

Scholl, B.J. & A.M. Leslie 1999. Modularity, development and 'theory of mind'. *Mind and language 14/1*, 131-153.

Sigman, M., Ungerer, J.M., Mundy, P. & Sherman, T. (1986) Social interactions of autistic, mentally retarded, and normal children and their caregivers, *Journal of Child Psychology and Psychiatry* 27, 647-656.

Trevarthen, C.B. (1977) Descriptive analyses of infant communicative behaviour, in Schaffer, H.R. (ed.), *Studies in Mother-Infant Interactions*, Londres, Academic Press.

Yun Chin, H. and Bernard-Opitz, V. (2000) Teaching Conversational Skills to Children with Autism: Effect on the Development of a Theory of Mind, *Journal of Autism and Developmental Disorders* 30/6, pp. 569-583.