

Psychologie cognitive et calcul quantique

« *Penser c'est calculer* », mais au fait....comment calcule-t-on ?

Introduction.

La psychologie cognitive est la sous-discipline de la psychologie qui traite tout particulièrement de la « cognition », ce dernier terme renvoyant, de façon très générale, à l'étude de la pensée et des processus mentaux impliqués dans la réalisation de tâches cognitives, telles que la perception, la mémorisation, la compréhension du langage, le raisonnement ou la prise de décision. Les psychologues cognitivistes cherchent donc à déterminer et à étudier les opérations mentales utilisées lorsque nous réalisons ces diverses tâches. Ils utilisent pour cela la méthode expérimentale initiée par l'approche behavioriste (Skinner 1969) mais, contrairement au refus affiché par cette dernière de prendre en compte les mécanismes intermédiaires entre stimulus et réponse (mécanismes qui sont relégués dans une « boîte noire », pour des raisons méthodologiques et conceptuelles), ils considèrent qu'il est nécessaire de les prendre en considération. Ils utilisent la notion d'« états mentaux », ne serait-ce que pour le rôle qu'ils jouent au sein du réseau de relations causales constitué par des entrées sensorielles, des sorties comportementales et d'autres états mentaux.

Cet article ne traite pas de la psychologie cognitive en tant que discipline, pour sa pratique et ses méthodes spécifiques, et n'aborde pas non plus les mécanismes neurobiologiques qui rendent possible la pensée -mécanismes qui ont inspiré l'approche connexionniste dont nous parlerons à la section 1. La question plus spécifiquement soulevée ici est celle de ses fondements conceptuels, et en particulier de *la thèse cognitiviste* qui lui a donné naissance et qui assimile la pensée à un calcul. Les processus mentaux qu'elle étudie sont-ils régis par la logique classique et le calcul classique des probabilités, comme le suggéraient Boole et Frege et comme cela semble acquis dans les modélisations qu'elle en propose en terme de traitement séquentiel d'information par un ordinateur (ou une machine de Turing) fonctionnant selon les algèbres de Boole ? Et en cas de réponse négative à cette première question, sa modélisation doit-elle faire appel aux logiques non-classiques –et si oui, lesquelles ?

Pour répondre à ces questions, nous commencerons (à la section 1) par rappeler brièvement les motivations et les enjeux qui ont permis l'éclosion du paradigme cognitiviste dans lequel

elle s'inscrit, un paradigme qu'aucune critique, d'ordre « technique » (connexionnisme) ou épistémologique (approche enactive de Varela), n'a vraiment réussi à ébranler. Nous expliquerons ensuite (à la section 2) pourquoi la notion *classique* de calcul sur laquelle elle s'est érigée pose problème. Ce point sera abordé à partir des travaux expérimentaux de Hampton dans le domaine de la linguistique (combinaison de concepts) et par ceux de Tversky et Shafir dans le domaine de la théorie de la décision. Nous soutiendrons alors, en accord avec Aerts et son équipe (Aerts et al. 2009) (Aerts et al. 2011), ainsi que Busemeyer et Bruza (2012), que les processus mentaux peuvent être modélisés de façon féconde selon une approche de type quantique capable de rendre compte des effets d'ordre, de contextualité et d'interférences auxquels ils donnent lieu. Et que c'est donc la notion de calcul qui fonde l'approche cognitive qui doit être re-précisée en ce sens.

1. Quid des « lois de la pensée » ?

Le projet inachevé de Leibniz était de créer une « *lingua characteristica universalis* », c'est à dire un système de signes idéographiques, en contact direct avec les idées qu'ils expriment, dans laquelle le raisonnement humain pourrait être traduit sous la forme d'un calcul. Afin de réaliser cette langue idéographique universelle dont la syntaxe ne se réduirait à aucune grammaire empirique, Leibniz entreprend de décomposer les idées complexes en idées simples jusqu'à atteindre ce qu'il appelle « l'alphabet des pensées humaines ».

Selon cette même idée, Boole propose dans « *Les lois de la pensée* » (Boole 1854) d'étudier « les lois fondamentales des opérations de l'esprit, les exprimer dans le langage symbolique du calcul... », ou encore « d'étudier les lois fondamentales des opérations de l'esprit par lesquelles s'effectue le raisonnement ». Il suggère que les opérations mentales sont régies par des lois semblables à celles qui gouvernent l'algèbre, ce qui l'amène à proposer une « algèbre des pensées humaines » fondée sur une combinatoire faisant intervenir seulement deux éléments (0 et 1) –et qui peut être considérée comme le fondement de la logique classique propositionnelle (qui est bivalente).

Enfin, toujours dans ce même ordre d'idées, mentionnons les travaux fondateurs de Frege en logique des prédicats qui, dans sa *Begriffsschrift* (Frege 1879), se propose de reprendre le projet inachevé de Leibniz. Son but est de construire une « langue formulaire de la pensée pure construite sur le modèle de l'arithmétique » dans laquelle on puisse exprimer son contenu. La création d'un tel symbolisme permettrait, en effet, de « fournir le critère le plus sûr de la validité d'une chaîne d'inférences et de nous permettre de remonter jusqu'à la source de tout ce qui y restait implicite », et ainsi de « peindre non pas les paroles, mais les pensées ».

Cette conception « algébrique » de la pensée et des processus cognitifs qui fait appel à un niveau symbolique de traitement de l'information se propose donc d'expliquer de cette façon comment nos états mentaux représentationnels (états cognitifs) et même intentionnels (désir, croyance, ...) peuvent déterminer nos comportements ou nos « conduites », selon le terme utilisé par P. Janet. Elle a reçu une impulsion décisive avec l'avènement de la théorie du traitement de l'information (Turing 1936) et de la cybernétique (Wiener 1948), ainsi que de la linguistique, avec la notion de grammaire générative proposée par Chomsky (1957). Le développement de cette approche « cognitive » de la cognition est, en fait, étroitement lié à la thèse de Church-Turing affirmant que la notion physique de la calculabilité peut être caractérisée à l'aide d'un ensemble de règles de calcul (celles définissant les fonctions récursives, par exemple) ainsi qu'à la découverte de McCulloch et Pitts (1943) selon laquelle les fonctions logiques peuvent être réalisées à l'aide de neurones formels à deux niveaux. Le couronnement de l'approche cognitive et, corrélativement, l'affirmation des bases conceptuelles de la psychologie cognitive, furent apportés par la réalisation des premiers programmes informatiques capables de simuler l'activité mentale -ce qu'a permis la construction des ordinateurs conçus selon l'architecture de von Neumann- et par les progrès fulgurants réalisés depuis, en particulier dans le domaine de l'intelligence artificielle¹ (Minsky 1988).

L'approche cognitive très rapidement exposée ci-dessus s'est donc construite à partir de l'idée que notre activité mentale pouvait être assimilée à un calcul *classique* et serait donc modélisable en terme de machine de Turing ou de programme d'ordinateur fonctionnant selon les algèbres de Boole. Elle s'est heurtée à deux types de critiques. La première concerne la limitation de sa capacité à réaliser certaines tâches cognitives d'importantes, comme la reconnaissance de formes ou l'apprentissage. En effet, un programme informatique peut être compromis si un seul symbole de ce programme vient à manquer et cette fragilité semble mal s'accorder avec la notion d'apprentissage par l'expérience qui caractérise ces tâches cognitives et encore moins avec l'idée de plasticité neuronale (Changeux 1983) selon laquelle les compétences cognitives du cerveau utilisées dans la réalisation de certaines tâches cognitives peuvent être restaurées (partiellement, du moins) en dépit d'une lésion locale affectant les

¹ L'intelligence artificielle conçoit des systèmes artificiels nous renseignant sur la manière dont les êtres vivants (humains et animaux) accomplissent des tâches intelligentes de difficulté variable -par exemple, des « systèmes experts » capables de comprendre une conversation ou de diagnostiquer une pathologie médicale, ce qui repose sur une représentation symbolique préalable de la situation et des actions à entreprendre pour réussir de telles tâches.

organes initialement sollicités dans ces tâches. En outre, la rigidité inhérente aux règles de calcul de ces programmes a pour conséquence que ces derniers ne fonctionnent plus dans le cas de la réalisation de tâches nécessitant une continuelle adaptation aux informations issues de l'environnement –comme c'est le cas, par exemple, pour la reconnaissance d'un visage dans un environnement bruité.

C'est en réaction à ces limitations du cognitivisme et même en opposition avec lui que s'est développée l'approche connexionniste de la cognition, proposant de prendre en compte les processus neurophysiologiques sous-jacents pour expliquer l'émergence de la pensée, en modélisant ces processus à l'aide de réseaux de neurones formels mimant le fonctionnement de réseaux de neurones biologiques. Cette approche s'est constituée à partir des travaux fondateurs de Mc Culloch et Pitts sur le neurone binaire brièvement mentionnés ci-dessus et ceux de Hebb (1949) qui a proposé une règle de renforcement des synapses susceptible de rendre compte de certains phénomènes d'apprentissage². Elle a franchi une étape décisive lorsque Rosenblatt (1958) a proposé le premier « Perceptron », modèle informatique d'un réseau de neurones formels répartis sur deux niveaux et capable de reconnaître les lettres de l'alphabet.

Le connexionnisme considère que le niveau symbolique n'est pas pertinent pour aborder l'étude des processus mentaux et propose de les traiter comme des processus émergents de réseaux de neurones fortement interconnectés qui permettraient le traitement massivement parallèle de l'information. Comme les neurones biologiques, les neurones formels sont susceptibles d'exciter ou d'inhiber les neurones auxquels ils sont connectés selon la règle de Hebb. Ils forment ainsi un système dynamique dont l'évolution est définie par l'introduction d'une « fonction d'activation » permettant de calculer l'état de chaque neurone en fonction de ceux des neurones dont il reçoit l'influence (via les « synapses » formelles). Ce système est donc capable de traiter de façon globale l'information qui lui est fournie à l'entrée et la tâche cognitive requise (apprentissage ou reconnaissance de formes, en particulier) ainsi réalisée par le réseau de neurones lorsque son état global est stabilisé.

L'approche cognitiviste sur laquelle s'est édifiée la psychologie cognitive a aussi été critiquée sur un plan *épistémologique* par Varela et ses collaborateurs dans la mesure où cette approche, tout comme l'approche connexionniste d'ailleurs, se fonde sur l'idée qu'il existerait *un monde pré-déterminé et pré-existant à la perception* que le cerveau « appréhenderait ». Selon cette idée (Varela 1988, p. 104): « le monde extérieur comporte des règles fixes ; il

² La règle de Hebb affirme que la connexion entre deux neurones se renforce lorsque ces neurones sont activés simultanément alors qu'elle s'affaiblit dans le cas contraire.

précède l'image qu'il projette sur le système cognitif dont la tâche est de le saisir -le monde- de manière appropriée (que cela soit au moyen de symboles ou d'états globaux) ».

En outre, le cognitivisme et le connexionnisme sont des approches « à la troisième personne », effectuée dans le langage de la science, où la dimension *phénoménale* de l'expérience est évacuée³. Contrairement à ces deux approches, l'approche enactive de Varela propose de réintégrer la dimension phénoménale dans le processus de la cognition, c'est à dire de tenir compte de l'expérience vécue tout en considérant les mécanismes neuronaux sous-jacents. En accord avec la proposition que fait Nagel dans son article « What it is like to be a bat ? » (Nagel 1974), Varela propose de rendre compte à la fois du vécu de l'expérience et de ce que l'on peut dire du fonctionnement cérébral correspondant, de ses corrélats neuronaux, en introduisant une notion de causalité réciproque entre ces deux domaines. La cognition apparaît donc comme un processus de co-naissance ou de co-émergence du sujet de la connaissance et de l'objet de la connaissance, ce qui a pour conséquence que le monde perçu est « enacté » en ce qu'il est construit au travers de notre expérience et non pas prédéterminé.

Cependant, malgré leur pertinence évidente, nous pouvons affirmer qu'aucune de ces deux critiques ne remet vraiment en question le projet cognitiviste de modéliser la pensée comme un calcul symbolique. En effet, si le cognitivisme modélise la pensée comme une manipulation de symboles selon des règles de syntaxe abstraites, le connexionnisme considère que la pensée « émerge » de l'activité de réseaux de neurones formels (qui sont censés reproduire l'activité du cerveau humain). Ces deux approches ont donc *des objets différents* : le connexionnisme modélise les *processus matériels*, voire biologiques, qui rendent (ou rendraient) possible la pensée, alors que le cognitivisme modélise de façon abstraite les *processus mentaux* en terme de calcul ou de programmes d'ordinateur. On peut dire, par conséquent, qu'au lieu de s'opposer elles *se complètent* dans la mesure où elles abordent l'étude des processus cognitifs à partir de *deux niveaux distincts de description* : le niveau matériel ou neuronal à partir duquel on pourrait expliquer leur possibilité (objet du connexionnisme) et le niveau symbolique qui décrit leurs propriétés et leur articulation de façon abstraite, chacune de ces descriptions s'avérant utile et féconde dans leurs domaines respectifs.

De même, on ne peut pas dire non plus que l'idée d'assimiler les processus cognitifs à un calcul symbolique soit vraiment remise en question par la critique qu'en fait l'approche enactive

³ Notons que cette évacuation de l'expérience subjective par les modèles actuels des processus mentaux et de la conscience qui ne parlent que de leurs corrélats neurophysiologiques et des processus informationnels qui leur sont associés a été critiquée par plusieurs auteurs dont Chalmers (1996) et elle présentée de façon concise dans la référence (Uzan 2013a, pp. 23-28).

de Varela. Ce qui est, à juste titre d'ailleurs, remis en question par cette critique d'ordre épistémologique est l'idée que la cognition puisse être réduite à l'enregistrement et au traitement de l'information obtenue d'un monde *pré-existant* à la perception que nous en avons et qu'elle ne prend donc pas en compte ses propriétés dynamiques, sociales et phénoménales. Cependant, concevoir pour ces raisons la cognition comme un *processus de co-naissance du sujet et de ses représentations* n'est pas incompatible avec le projet de modélisation abstraite de ces mêmes représentations et des opérations mentales qu'elles mettent en jeu, même si ces opérations ne sont pas indépendantes du contexte dans lequel elles sont effectuées (voir ci-après et la section 2) et dépendent aussi, comme l'a montré Piaget (1928), du stade d'évolution psychologique du sujet.

Par contre, et c'est là, à notre avis, son véritable point faible, l'approche cognitiviste à partir de laquelle s'est édifiée la psychologie cognitive suppose que les processus cognitifs pourraient être assimilés à un « *calcul* » *classique*, c'est à dire qu'ils pourraient être modélisés en utilisant les règles de syntaxe de la logique classique, le calcul classique des probabilités (axiomatisé par Kolmogorov) et les algèbres de Boole. Il prend pour fondement inconditionnel le calcul *classique* tel qu'il est implémenté par la machine de Turing ou l'ordinateur classique. Ce qui suppose, en outre, d'accepter un *principe de réalisme des états mentaux* qui est souvent passé sous silence. Ce principe stipule qu'il y aurait toujours un sens à affirmer que nos états mentaux, qui sont transformés selon les algorithmes classiques définissant ces programmes, sont *parfaitement déterminés et définissables* à chaque instant (même si nous ne les connaissons pas). En effet, selon le principe de bivalence de la logique classique qui est présupposée dans l'approche cognitiviste, un énoncé affirmant qu'un sujet se trouve dans un certain état mental est soit vrai soit faux, il ne peut être « indéterminé » -contrairement à ce qui peut être le cas en réalité, comme nous le verrons à la section suivante.

Cette approche de la cognition (et de l'intentionnalité pour certains chercheurs se proposant de « naturaliser » cette dernière) qui est fondée sur le paradigme du traitement classique de l'information pourrait ainsi, pour reprendre en le complétant un slogan formulé dans de nombreux ouvrages de sciences cognitives, être caractérisée par la devise « Penser, c'est calculer *selon le calcul classique* ».

Cependant, il semblerait que cette conception classique du calcul à partir de laquelle s'est construit le cognitivisme et, dans la foulée, la psychologie cognitive ne puisse résister à l'épreuve de l'expérience et doit être révisée. Par exemple, nos raisonnements peuvent être inductifs plutôt que déductifs, leurs conclusions étant évaluées avec un degré de vérité sous l'hypothèse que ses prémisses sont vraies, ou dans certains cas la prise en compte d'une

nouvelle information peut réduire ce que nous savions d'une situation –ce qui viole la propriété de monotonie de la logique classique. Pour résoudre ces difficultés, il a été fait appel à des logiques non-classiques, comme les *logiques floues* (Dernancourt 2011) qui capturent l'idée selon laquelle nos raisonnements peuvent n'être qu'« approchés » et introduisent une notion de « degré de vérité » d'un énoncé –comme le font aussi les logiques probabilistes-, ou les *logiques non-monotones* (Lecomte 2005) qui rendent compte du processus de restriction des connaissances auquel peut donner lieu la prise en compte d'une nouvelle information –une propriété caractéristique de nos raisonnements quotidiens. Nous ne remettons pas en question la fécondité de ces logiques non-classiques (ni d'autres logiques non classiques d'ailleurs, comme la logique linéaire ou les logiques épistémiques et temporelles) qui sont ou pourraient être utilisées avec succès pour modéliser nos processus mentaux⁴. Cependant, nous montrerons à la section suivante, en nous référant à des résultats expérimentaux qui sont représentatifs de notre façon de raisonner, de combiner des concepts ou de prendre des décisions, que certaines propriétés importantes des processus cognitifs (effets d'ordre, de contextualité et interférences), propriétés qui remettent en question l'interprétation des connecteurs qui est faite par la logique classique et ses extensions (comme les logiques modales et la logique floue), semblent relever d'une approche quantique.⁵

2. Processus cognitifs et théorie quantique.

D. Aerts et son équipe du centre Leo Apostel de Bruxelles (Aerts 2009) (Aerts et al. 2011) ainsi que Busemeyer et Bruza (2012) ont proposé de modéliser les processus cognitifs selon une approche quantique afin de rendre compte des *effets d'ordre*, de *contextualité* et d'*interférences* qui leur sont inhérents. Cette approche qui utilise le calcul *quantique* des probabilités s'inscrit dans le cadre d'une théorie quantique « généralisée », où toute référence a priori au monde physique a été éliminée, et qui a été initialement proposée par Atmanspacher, Römer, Filk et Walach (Atmanspacher et al. 2002, 2009) (Walach and von Stillfried 2010). Cette

⁴ Notons que ces deux logiques non classiques sont déjà utilisées avec succès dans le domaine de l'intelligence artificielle, par exemple en médecine pour l'aide au diagnostique et à la décision ou en robotique.

⁵ Ce qui signifie que, sans amoindrir les promesses des logiques non classiques et, en particulier, des logiques non-monotones à prendre en compte ces propriétés, le calcul quantique semble se prêter tout « naturellement » à cet exercice puisque ces propriétés sont inscrites dans sa structure même par la définition qu'il donne des connecteurs logiques.

théorie a été utilisée selon une version légèrement différente⁶ dans plusieurs articles parus récemment dans *Implications Philosophiques* et dans *Kronoscope* pour expliquer les corrélations psychophysiques en terme de relation d'intrication, c'est à dire sans faire appel à un concept de causalité efficiente (Uzan 2012a, b) –une interprétation qui devrait bientôt recevoir une confirmation expérimentale (Uzan 2014a). L'approche quantique du phénomène psychosomatique, qui a permis l'élaboration de modèles explicatifs et prédictifs (Uzan 2013b), ne sera donc pas reprise ici. La thèse que nous soutiendrons en accord avec les chercheurs mentionnés ci-dessus est que cette approche de type « quantique » s'avère aussi extrêmement féconde pour modéliser les processus cognitifs. Nous allons expliquer ce point en l'illustrant à l'aide d'exemples relevant des domaines de la sémantique (combinaison de concepts), du raisonnement et de la théorie de la décision.

2.a La combinaison de concepts.

Le calcul classique obéit au principe de vérifonctionnalité selon lequel la valeur de vérité d'un énoncé complexe est déterminée par la valeur de vérité des énoncés plus simple qui le composent en appliquant un nombre fini de règles qui interprètent les connecteurs utilisés dans cet énoncé. De façon analogue, l'approche computationnelle de la cognition adopte la thèse de compositionnalité de la *signification*. Selon cette thèse, qui a été soutenue par Pelletier (1994) et par Fodor (1994), la signification d'un énoncé serait exclusivement déterminée par celles de ses composants et la façon dont ils se combinent. Cependant, l'expérience quotidienne semble nous montrer que la thèse de compositionnalité de la signification est remise en question : la signification que nous attribuons à une expression linguistique dépend de la phrase ou du texte dans lequel elle est utilisée et cette signification se définit ainsi par sa contribution à ce contexte linguistique global. Cette propriété de *contextualité* du langage naturel a été précisée et généralisée par plusieurs auteurs dont Searle (1992) et Travis (1997), dans le prolongement des travaux fondateurs d'Austin (1961) et de Wittgenstein (1953). Ce dernier soutient dans ses *Investigations Philosophiques* que la signification d'un énoncé ne peut se déterminer par la seule application de règles logiques mais qu'elle est définie par *ses règles d'usage* à l'intérieur de « jeux de langages » ou modes de communication relatifs aux pratiques sociales et culturelles –ce qui fait que la signification d'un énoncé dépend du contexte linguistique, mais aussi social et culturel où il est prononcé ou utilisé.

⁶ Nous avons, en effet, suggéré de ne pas affaiblir la structure mathématique sous-jacente de cette théorie (algèbre C* d'observables et représentation des états dans un espace de Hilbert) .

Une étude expérimentale récente conduite par Hampton (1988a, b) dans le prolongement des travaux de Osherson et Smith (1981) sur « l'effet Guppy »⁷ a permis de quantifier cette propriété de contextualité du langage dans un cas très simple, en étudiant la façon dont nous assignons une signification à un concept obtenu en composant deux concepts. Les travaux de Hampton montrent que les opérations mentales relatives à la conjonction et la disjonction de deux concepts *ne sont pas régies par le calcul des probabilités classiques*. Pour établir ce résultat, Hampton commence par définir une sémantique extensionnelle pour chacun des concepts utilisés: les sujets testés doivent choisir, sur une échelle de nombres entiers croissants allant de -3 à +3, le degré d'appartenance d'un item à ce concept (ou dans quelle mesure il peut être considéré comme référent pour ce concept), ce qui définit, pour un concept noté « c », un ensemble $E(c)$ pondéré d'items qui peuvent lui être assignés comme référents possibles. Les fréquences obtenues sont rapportées à un intervalle continu [0, 1] et définissent ainsi une distribution de probabilités associées à chacun des items présentés pour un concept donné. Pour un item donné, ce test est effectué pour un concept A, pour un concept B et pour le concept obtenu en effectuant leur conjonction (A et B), ou leur disjonction (A ou B). Hampton montre alors que les probabilités pour qu'un item donné appartienne au concept A, au concept B puis à leur conjonction A et B *ne sont pas reliées selon règles élémentaires du calcul classique des probabilités*, et il en est de même pour la disjonction de ces concepts.

Par exemple, Hampton (1988a, p. 31) mesure sur un groupe de sujets qui n'étaient pas initialement informés de la nature de l'expérience (et que l'on peut donc considéré comme choisi de façon aléatoire) le degré d'appartenance de l'item « garlich » (ail) aux concepts « food » (nourriture), au concept « plant » (plante) et à leur conjonction « food plant » (plante comestible). La probabilité $P(f)$ avec laquelle « garlich » est considéré comme référent du concept « food » est 0,95; celle, notée $P(p)$, avec laquelle « garlich » est considéré comme référent du concept « plant » est 0,70; alors que celle, notée $P(f \text{ et } p)$, avec laquelle « garlich » est considéré comme référent du concept « food plant » est 0,79. Le fait que $P(f \text{ et } p)$ soit plus grand que le minimum des deux valeurs $P(f)$ et $P(p)$, qui est ici $P(p) = 0,70$, est *incompatible* avec les règles du calcul classique des probabilités. En effet, selon la loi de croissance de la probabilité, qui est une conséquence des axiomes du calcul classique des probabilités énoncés par Kolmogorov, la probabilité $P(f \text{ et } p)$ ne peut excéder ni $P(f)$ ni $P(p)$ puisque l'intersection $E(f) \cap E(p)$ des extensions des concepts « food » et « plant » est incluse dans $E(f)$ et dans $E(p)$.

⁷ L'effet « Guppy » désigne, en fait, un paradigme historique de l'effet contextuel du langage, pour lequel l'appartenance du référent « Guppy » au concept composé « Pet-Fish » ne se réduit pas à la simple combinaison de celles des concepts « Pet » et « Fish » (Aerts 2011).

Par contre, ce phénomène de « sur-extension » des concepts, appelé ainsi par Hampton, peut s'expliquer en supposant que par la nature de l'item proposé (« garlich »), qui est généralement conçu tout d'abord comme un aliment par les sujets interrogés, la signification du concept « plant » est influencée ou dominée par le concept « food » qui joue donc le rôle de *contexte* dans lequel est interprété le concept composé « food plant ».

Ces résultats ont été analysés par Aerts et son équipe (2011) qui ont suggéré que la propriété de contextualité mise en jeu dans cette opération de composition conceptuelle pouvait être expliquée et modélisée selon une approche quantique généralisée. Aerts et Gabora (2005) ont, en effet, proposé un modèle quantique très élaboré afin de rendre compte de façon précise du résultat précédent (et d'autres résultats similaires). Cependant, par souci pédagogique, nous allons présenter ici un modèle simplifié qui en saisit les idées essentielles et permet d'expliquer l'anomalie observée.

La signification d'un concept est, comme dans les travaux de Hampton, définie selon une sémantique extensionnelle par l'ensemble de ses référents possibles qui sont conçus comme ses « états » possibles. Ces états sont représentés par des vecteurs unitaires dans un espace de tous les états possibles –en fait, un espace de Hilbert- où ils forment une base orthonormée du sous-espace relatif à ce concept. Le concept A est ainsi représenté dans ce modèle par un vecteur normé qui peut s'écrire comme la combinaison linéaire suivante :

$$|A\rangle = \sum_i \alpha_i |A_i\rangle,$$

où les vecteurs $|A_i\rangle$, qui représentent ses référents possibles, sont affectés des composantes complexes α_i obéissant à la condition de normalisation $\sum_i |\alpha_i|^2 = 1$. Le fait de représenter un concept par une *superposition d'états possibles* et non pas seulement par l'un d'eux traduit l'idée selon laquelle, contrairement à ce qui est affirmé par le principe de réalisme des états mentaux du calcul classique (voir section 1), on ne suppose pas ici que l'état mental d'un participant à ce test (qui est décrit ici par la signification qu'il attribue à ce concept) est toujours déterminé de façon unique et non ambiguë.

La probabilité pour que l'item X appartienne au concept A se calcule, dans ce modèle, en évaluant la probabilité que le vecteur unitaire $|X\rangle$ associé à X soit l'un des référents possibles de A, c'est à dire l'un des vecteurs $|A_i\rangle$ de la décomposition ci-dessus –ou, autrement dit, c'est la probabilité de trouver le concept A dans l'état $|X\rangle$. Cette dernière probabilité, qui peut se calculer en utilisant la règle de Born de la théorie quantique⁸, est égale au carré du module du

⁸ Notons que contrairement aux apparences (justifiée uniquement par une approche historique de la physique quantique), cette règle n'a rien de spécifiquement physique puisque, comme l'a

vecteur $\Pi(A)|X\rangle$ obtenu en projetant orthogonalement $|X\rangle$ sur $|A\rangle$, ce qui s'écrit $P(A) = \langle X|\Pi(A)|X\rangle$. Et on peut calculer de façon similaire la probabilité $P(B)$ pour que l'item X appartienne au concept B , par $P(B) = \langle X|\Pi(B)|X\rangle$.

Lorsqu'un sujet évalue la probabilité que X appartienne à la conjonction (A et B), on peut supposer qu'il commencera en général par évaluer cette probabilité pour l'un des deux concepts, disons A , avant l'autre –cet ordre s'expliquant par le fait que la probabilité d'appartenance de X est, *pour ce sujet*, plus grande pour A que pour B et qu'il a donc une préférence à attribuer X à A plutôt qu'à B . L'estimation de cette probabilité se fera donc en deux temps (même si ces instants sont très rapprochés) : l'action du projecteur $\Pi(A)$, qui donne pour résultat $\Pi(A)|X\rangle$, *établit le contexte* dans lequel le sujet va évaluer l'appartenance de X au concept (A et B), ce qui se traduit formellement par l'application du projecteur $\Pi(B)$ à ce premier résultat pour donner $\Pi(B)\Pi(A)|X\rangle$. La probabilité correspondante est donc ici :

$$P(A \text{ et } B) = P(A \text{ puis } B) = \langle X|\Pi(B)\Pi(A)|X\rangle.$$

Si, comme nous l'avons supposé, $P(A) > P(B)$, ce que traduit la préférence du sujet à attribuer X à A en premier, il est alors possible de montrer que la probabilité $P(A \text{ et } B)$ est plus grande que le minimum de $P(A)$ et de $P(B)$, qui est ici $P(B)$ -ce qui peut donc expliquer l'anomalie observée (voir ci-après pour plus de détails).

La figure 1 ci-dessous permet de visualiser ce calcul et de vérifier la violation de la règle de croissance du calcul classique des probabilités. Sur ce schéma où les vecteurs sont supposés être de longueur unité, le vecteur $|X\rangle$ représentatif de l'item testé se projette sur $|A\rangle$ puis après sur $|B\rangle$, ce qui donne le vecteur noté $|Y\rangle$. La probabilité $P(A \text{ et } B)$, qui est égale dans ce cas à $P(A \text{ puis } B)$, est la longueur du vecteur obtenu en projetant $|Y\rangle$ sur $|X\rangle$ -voir la construction en bleu. Quant à la probabilité $P(B)$, elle est obtenue en projetant $|X\rangle$ directement sur $|B\rangle$ puis en projetant le vecteur obtenu, noté $|Z\rangle$, sur $|X\rangle$ -construction en vert. Ce schéma montre que si $P(A) > P(B)$, l'angle formé entre le vecteur $|X\rangle$ et le vecteur $|A\rangle$, qui est noté « a » sur cette figure, est plus petit que l'angle, noté « b », formé entre $|X\rangle$ et le vecteur $|B\rangle$, ce qui a pour conséquence que $P(A \text{ et } B) > P(B)$.

montré Gleason, elle spécifie l'unique mesure de probabilité pouvant être définie sur les sous-espaces d'un espace de Hilbert (Gleason 1957).

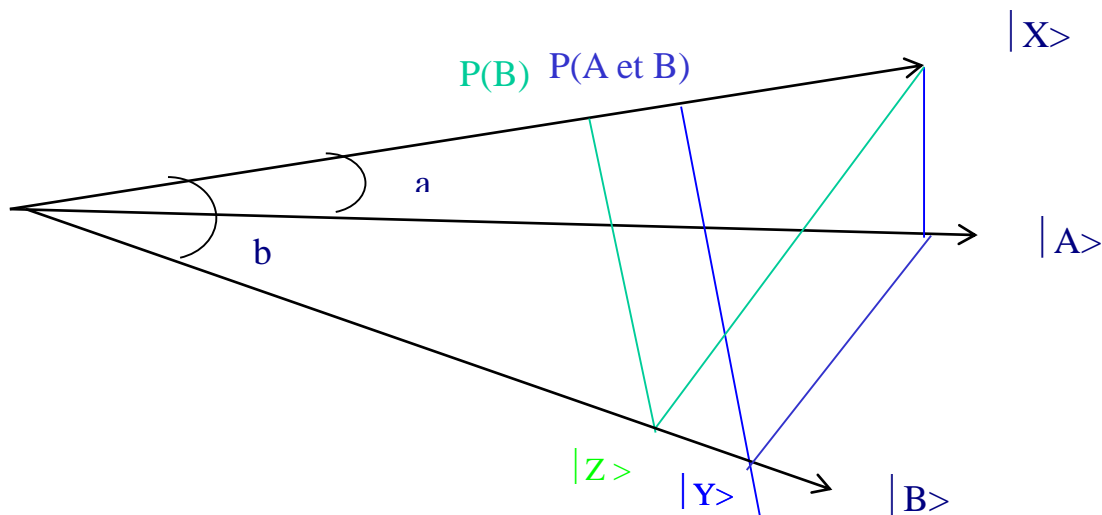


Figure 1. Modèle géométrique pour le calcul des probabilités $P(A \text{ puis } B)$ et $P(B)$.

De façon plus précise, ce modèle permet de retrouver avec une bonne approximation les résultats expérimentaux de Hampton rapportés dans son article (Hampton 1988a), en calculant tout d'abord la valeur des angles a et b à partir des probabilités $P(A)$ et $P(B)$ *expérimentales*. Ces probabilités sont, en effet, reliées aux angles a et b de la façon suivante (voir la figure 1) :

$$P(A) = \langle X | \Pi(A) | X \rangle = \cos^2 a ; P(B) = \langle X | \Pi(B) | X \rangle = \cos^2 b,$$

Puis, dans un deuxième temps, on calcule $P(A \text{ et } B)$ à partir de ces angles :

$$P(A \text{ et } B) = \langle X | \Pi(B) \Pi(A) | X \rangle = \cos a \cos b \cos (b-a).$$

Pour l'exemple considéré ci-dessus, les valeurs des angles a et b calculées à partir des valeurs expérimentales rapportées ci-dessus $P(A) = 0,95$ et $P(B) = 0,70$ sont $a = 12,9^\circ$ et $b = 33,2^\circ$, et donc on peut calculer $P(A \text{ puis } B) = 0,765$. Nous retrouvons donc avec une bonne approximation la valeur expérimentale rapportée par Hampton (et ci-dessus), qui est 0,79.

2.b. Raisonnements « fallacieux » et paradoxes de la théorie classique de la décision.

Tversky et Shafir (1992) ont montré sur des bases expérimentales une violation significative du principe de cohérence des choix possibles de la théorie classique de la décision qui traduit, dans ce domaine, la validité de la loi des probabilités totales du calcul classique des probabilités. Cette violation donne lieu aux paradoxes d'Allais et d'Ellsberg relatif au comportement des agents économiques (Ellsberg 1961).

Le principe de cohérence des choix possibles, appelé dans la littérature « sure thing principle », peut se formuler de la façon suivante (Savage 1954):

« Si A est préféré à B lorsque X réalisé et si A est préféré à B lorsque non X réalisé, alors A est préféré à B même si on ne sait pas si X est ou non réalisé. »

Ce principe a été testé par Tversky and Shafir dans l'expérience suivante : les participants se voient (réellement) proposer de jouer à un jeu où la chance de gagner 200 \$ est la même que celle de perdre 100 \$. Ce jeu est proposé une première fois aux participants et on leur demande s'ils veulent jouer une seconde fois dans les trois cas suivants (la population étant alors répartie en trois groupes de même taille) : (a) ils savent qu'ils ont gagné la première fois, (b) ils savent qu'ils ont perdu la première fois et (c) *ils ne savent pas* s'ils ont gagné ou perdu la première fois. Les résultats sont les suivants : 69% des participants du groupe (a) décident de rejouer, 59% des participants du groupe (b) décident de rejouer, alors que seulement 36% des participants du groupe (c), dont les participants ne savent pas s'ils ont gagné ou perdu, décident de rejouer. Ce résultat est en désaccord avec le principe de la chose sûre mentionné ci-dessus puisque si ce dernier s'appliquait, on obtiendrait pour le groupe (c) une probabilité égale à la demi-somme des probabilités mesurées, respectivement, pour le groupe (a) et pour le groupe (b). En effet, dans la mesure où les deux événements possibles « avoir gagné » et « avoir perdu » équiprobables sont exclusifs, la loi des probabilités totales du calcul classique des probabilités nous dit que :

$$\begin{aligned}
 P(\text{rejouer}) &= P(\text{d'avoir gagné la première fois et de rejouer}) \\
 &\quad + P(\text{d'avoir perdu la première fois et de rejouer}) \\
 &= (1/2) [P(\text{de rejouer sachant qu'on a gagné la première fois}) \\
 &\quad + P(\text{rejouer sachant qu'on a perdu la première fois})].
 \end{aligned}$$

Cette violation du calcul classique des probabilités, qui semble à première vue indiquer que les êtres humains seraient enclins à faire des raisonnements « fallacieux », constitue, en fait, un exemple typique de la façon dont nous « raisonnons » et prenons des décisions. Aerts et son équipe (Aerts and Sozzo 2013) ainsi que, selon une approche légèrement différente, Bussemeyer et Bruza (2012, chap. 9) ont montré que ce paradoxe peut être résolu si on représente les états et processus mentaux mis en jeu dans un cadre quantique généralisé. En effet, dans un tel cadre il est possible de montrer que lorsque la décision est prise en situation incertaine, la différence entre la valeur de la probabilité classique (de préférer A à B) et le résultat expérimental peut se calculer à partir du *terme d'interférence* entre les états mentaux possibles d'un participant (définis ici par ce qu'il sait ou non du résultat du jeu et/ou par son intention de jouer ou non). Plus précisément, définissons les deux suites possibles d'états mentaux h_1 et h_2 (ou « histoires » possibles) conduisant un participant à la décision de rejouer. L'histoire h_1 est définie par la suite d'états (I,G,J) et l'histoire h_2 est définie par la suite d'états (I, \bar{I} G,J), où :

I est l'état mental « initial » du sujet où il ne sait pas encore s'il va gagner ou perdre mais où il prend la décision de jouer une première fois,

G est son état mental lorsqu'il a appris qu'il a gagné la première fois

$\neg G$ est son état mental lorsqu'il a appris qu'il a perdu la première fois

J est son état mental lorsqu'il prend la décision de rejouer compte tenu du premier résultat.

Ces histoires possibles peuvent être schématisées de la façon suivante :

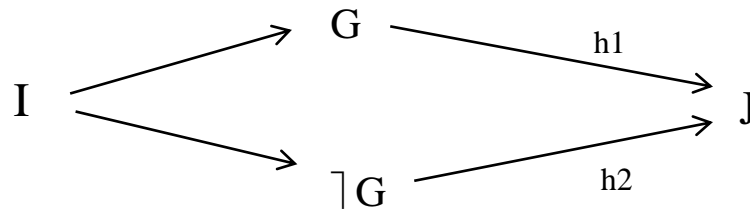


Figure 2. Les deux histoires possibles menant à la décision de rejouer.

Le calcul classique des probabilités (loi des probabilités totales) nous dit que :

$$P(J) = (1/2) [P(J/G) + P(J/\neg G)],$$

qui est, comme nous venons de le voir, *en contradiction* avec les données expérimentales. Par contre, ce paradoxe peut être résolu *si nous supposons ici aussi que*, contrairement à ce qu'affirme le principe de réalisme des états mentaux du calcul classique, *le participant qui ne sait s'il a ou non gagné la première fois se trouve dans un état de superposition des états G et $\neg G$* . Soit $|g\rangle$ cet état, qui est représenté, comme dans l'exemple précédent, par un vecteur unitaire de l'espace de configuration des états mentaux (espace de Hilbert) et qui s'écrit comme la superposition des états G et $\neg G$:

$$|g\rangle = (1/\sqrt{2}) (|G\rangle + |\neg G\rangle),$$

où $|G\rangle$ et $|\neg G\rangle$ sont les vecteurs unitaires représentatifs des états G et $\neg G$, alors que le coefficient de normalisation $(1/\sqrt{2})$ est calculé en tenant compte de l'égalité des probabilités de gagner et de perdre la première fois. En appliquant maintenant la règle de Born à la transition de $|g\rangle$ à l'état $|J\rangle$ décrivant la décision de rejouer (et compte tenu du fait que la probabilité que le sujet parvienne à l'état $|g\rangle$ est égale à 1), on peut calculer la valeur de la probabilité de rejouer selon ce modèle quantique :

$$\begin{aligned} P'(J) &= \langle g | \Pi(J) | g \rangle \\ &= (1/2) [\langle G | \Pi(J) | G \rangle + \langle \neg G | \Pi(J) | \neg G \rangle + \langle G | \Pi(J) | \neg G \rangle + \langle \neg G | \Pi(J) | G \rangle] \\ &= (1/2) [P(J/G) + P(J/\neg G)] + \| \langle \neg G | \Pi(J) | G \rangle \| \cos w, \end{aligned}$$

où à la demi-somme des probabilités de rejouer sachant que l'on a, respectivement, gagné ou perdu la première fois (résultat du modèle classique) s'ajoute le *terme d'interférences* :

$$\text{Int} = \|\langle \bar{J} | G | \Pi(J) | G \rangle\| \cos w,$$

où w est l'argument du nombre complexe $\langle \bar{J} | G | \Pi(J) | G \rangle$. C'est la présence de ce terme qui explique l'écart entre les résultats expérimentaux reportés ci-dessus et ce que nous devrions trouver en appliquant la loi des probabilités totales. Ce terme d'interférence Int peut s'évaluer en commençant par calculer son carré en fonction des probabilités $P(J/G)$ et $P(J/\bar{J}G)$:

$$\text{Int}^2 = \|\langle \bar{J} | G | J \rangle \langle J | G \rangle\|^2 \cos^2 w = P(J/G) \cdot P(J/\bar{J}G) \cos^2 w.$$

Pour $w = 116^\circ$ on calcule que $\text{Int} = -0,28$ et donc $P'(J) = 0,36$ qui est conforme au résultat expérimental. On peut dire que l'angle w qui détermine la valeur du terme d'interférences mesure donc le facteur psychologique caractérisant la sensibilité (moyenne) des participants à l'information qui leur apprend (ou non) qu'ils ont gagné ou perdu au premier tirage. Autrement dit, w mesure leur aptitude à modifier leur décision de rejouer à la réception de cette information.

Notons enfin que Pothos et Busemeyer (2009) ont aussi présenté un modèle de type quantique afin d'expliquer l'écart entre la probabilité $P(J)$ obtenue par le calcul classique et le résultat expérimental rapporté ci-dessus⁹. Ce modèle, dont l'ingrédient essentiel est, comme pour le modèle simple présenté ci-dessus, la prise en compte du phénomène d'interférences entre les états mentaux et donc entre les chemins d'évolution possibles h_1 et h_2 , introduit, en outre, une notion d'évolution temporelle des états mentaux qui obéirait à une équation de type de l'équation de Schrödinger en physique quantique. Comme dans le modèle présenté ci-dessus, le modèle de Pothos et Busemeyer retrouve les résultats expérimentaux grâce à un choix approprié des valeurs des paramètres caractérisant la réaction (psychologique) des joueurs aux différents événements « apprendre qu'on a gagné », « apprendre qu'on a perdu » ou « ne pas savoir si on a gagné ou perdu ». Formellement, il s'agit de définir de façon appropriée les éléments de la matrice d'évolution régissant la transition d'un état mental à un autre et qui nous permet de calculer le terme d'interférences. Ce modèle ne sera pas plus développé ici, pour des raisons pédagogiques et aussi pour une question de place.

⁹ En fait, cet article traite aussi du paradoxe soulevé par le « dilemme du prisonnier » dont la résolution dans un cadre quantique est tout à fait similaire à celle présenté ici et montre que ces deux paradoxes ne peuvent pas trouver de solution satisfaisante dans le cadre d'une approche classique –en utilisant un modèle stochastique approprié à la situation (chaîne de Markov).

Conclusion.

Nous avons commencé par effectuer une brève description des concepts et des avancées techniques qui ont permis l'avènement de la psychologie cognitive. Cette branche de la psychologie qui se propose d'expliquer nos comportements en rendant compte des processus mentaux mis en jeu lorsque nous pensons et prenons des décisions est fondée sur l'approche cognitiviste de la cognition. Cette dernière approche assimile les processus mentaux à un calcul *classique*, régi par les règles de la logique classique et du calcul classique des probabilités et présuppose, en outre, la validité d'un principe de réalisme des états mentaux. Cependant, si aucune critique d'ordre « technique » (connexionisme) ou conceptuelle (approche énaactive) n'a pu remettre en question la fécondité de l'approche cognitiviste, c'est l'expérience qui nous montre que les processus mentaux ne peuvent généralement pas être assimilés à un calcul *classique*. Comme nous l'avons montré en nous référant aux travaux expérimentaux et théoriques de Hampton, de Aerts, de Busemeyer et de leurs équipes respectives, ces processus peuvent être expliqués et modélisés de façon féconde dans le cadre d'une théorie quantique généralisée où il devient possible de rendre compte des effets d'ordre, de contexte et d'interférences auxquels donnent lieu nos raisonnements, nos prises de décision et la façon dont nous combinons des concepts. Cette approche quantique de la cognition donne lieu à des modèles qui sont régis par une notion de calcul qui est plus générale que le calcul classique, plus « générale » en ce qu'elle se réduit au calcul classique lorsque nous négligeons ces effets typiquement quantiques : il s'agit toujours d'un calcul, mais d'un *calcul quantique*.

Bibliographie

- Aerts D. 2009: Quantum Structure in Cognition. ArXiv:0805.3850v2 [math-ph] 12 Mar 2009.
- Aerts D. and Gabora L. 2005: A Theory of Concepts and Their Combinations (II). *Kybernetes* 34, pp. 192-221.
- Aerts D. and Sozzo S. 2013: Quantum Structure in Economics: The Ellsberg Paradox. ArXiv:1301.0751 v1 [physics.soc-ph] 4 Janv 2013.
- Aerts, D., Gabora L., Sozzo S. and Veloz T. 2011: Quantum structure in cognition.: Fundamentals and Applications. ArXiv: 1104.3344v1 [cs.AI].

Atmanspacher H., Filk T. et Römer H. 2009: Théorie quantique faible : cadre formel et applications, in M. Bitbol (ed.), *Théorie quantique et sciences humaines*, CNRS Editions.

Austin J.L. 1961: *Philosophical Papers*, Oxford, Clarendon Press, 1979, 3e éd. Barsalou.

Boole G. 1854-1992: Les lois de la pensée. Traduit de l'anglais : An Investigation on the Laws of Thought, on which are founded the mathematical Theories of Logic and Probabilities (1854), par Souleymane Bachir Diagne. Paris : Vrin.

Busemeyer J. R. and Bruza P ; D. 2012: *Quantum Models of Cognition and Decision*. Cambridge University Press.

Chalmers D. 1996: *The Conscious Mind*. Oxford: Oxford University Press.

Changeux J. P. 1983 : *L'Homme neuronal*. Fayard Paris.

Chomsky N. 1957: *Syntactic Structures*, The Hague/Paris: Mouton.

Dernoncourt F. 2011 : La Logique Floue: entre raisonnement humain et intelligence artificielle. Mémoire de CogMaster.
https://www.academia.edu/1053137/La_Logique_Floue_entre_raisonnement_humain_et_intelligence_artificielle_FR_

Ellsberg D. 1961 : Risk, Ambiguity and the Savage Axioms. *Quarterly Journal of Economics* 75 (4), 643-669.

Frege G. 1879: *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle: Louis Nebert. Traduction anglaise: *Concept Script, a formal language of pure thought modelled upon that of arithmetic*, par S. Bauer-Mengelberg in Jean Van Heijenoort, ed., 1967. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*. Harvard University Press.

Gleason AM (1957) Measures on the closed subspaces of a hilbert space. *J. Math. Mech.* 6, pp 885–893

Hampton, J. A. 1988a : Overextension of conjunctive concepts: Evidence for a unitary model for concept typicality and class inclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 12-32.

Hampton, J. A. 1988b : Disjunction of natural concepts. *Memory & Cognition*, 16, 579-591

Hebb, D.O. (1949). *The Organization of Behavior*. New York: Wiley & Sons.

McCulloch W. S. and Pitts W. 1943: A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5: 115—133.

Minsky M. 1988 : *The Society of Mind* Simon and Schuster, New York.

Nagel T. 1974: What Is It Like to Be a Bat?, *The Philosophical Review*, Vol. 83, No. 4, pp. 435-450.

Piaget J. 1928: *The Child's Conception of the World*. London: Routledge and Kegan Paul.

Lecomte A. 2005 : Y a-t-il une logique de la classification ? in Polity Y., Henneron G. et Palermi R. (ed.): *L'organisation des connaissances: approches conceptuelles*, pp 21-35. L'Harmattan.

Pothos M. and et Busemeyer R. 2009 : A quantum probability explanation for violations of 'rational' decision theory. *Proceedings of the Royal Society B*, published online 25 March 2009.

Rosenblatt F. 1957: The perceptron –a perceiving and recognizing automaton. Report 85-460-1, Cornell. Aeronautical Laboratory.

Skinner B. F. 1969. *Contingencies of reinforcement : a theoretical analysis*. New York: Meredith Corporation. Traduction Française (1971): *L'analyse expérimentale du comportement*. Bruxelles : Mardaga.

Turing A. 1936 : On Computable Numbers with an Application to the Entscheidungsproblem. *Proc. London Math. Soc.*, 2ième Série, vol. 42 , 230-265.

Uzan P. 2012a Une approche quantique du problème corps-esprit (1) et (2). *Implications Philosophiques*. Octobre 2012.

Uzan P. 2012b A Quantum Approach to the Psychosomatic Phenomenon: Co-Emergence and Time-Entanglement of Mind and Matter. *Kronoscope, the journal of the International Society for the Study of Time* n°2, vol. 12, 219-244.

Uzan P. 2013a *Conscience et Physique Quantique*. Vrin, collection Mathesis.

Uzan P. 2013b The Quantum-Like Approach of Psychosomatic Phenomena in Application. *Axiomathes*, published online 23 december 2013.

Uzan P. 2014a. Deciding the question about the nature of psychophysical correlations. *Mind and Matter*. À paraître (juin 2014)

Varela F. 1988. *Invitation aux sciences cognitives*, Edition Seuil.

Walach H. and Stillfried (von) N. 2010: Generalised Quantum Theory—Basic Idea and General Intuition: A Background Story and Overview. *Axiomathes*, Special Issue 2.

Wiener N. 1948. *Cybernetics, or Control and Communication in the Animal and the Machine*. Paris: Hermann & Cie, MIT Press (Cambridge, Mass) et Wiley (New York).

Wittgenstein L. 1953 : *Investigations philosophiques*. Ed. G.E.M Anscombe et Rush Rhees. Traduction française 1961, rééd. Gallimard, coll. « Tel », 1986.