

La neuroéconomie pour comprendre les interactions sociales

Pierre Livet
Professeuse émérite Université d'Aix-Marseille CEPERC

Christian Schmidt
Professeuse émérite Université Paris Dauphine
Président de l'European Neuroeconomics Association

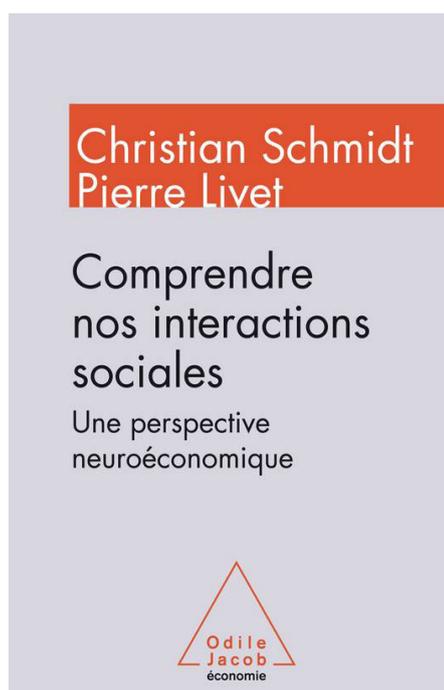
notes prises par **Hervé Dumez**
i3-CRG École polytechnique CNRS Paris-Saclay

LE 13 MARS 2015,
LE SÉMINAIRE
AEGIS A ACCUEILLI
PIERRE LIVET ET
CHRISTIAN SCHMIDT
À PROPOS DE LEUR
LIVRE CONSACRÉ À
UNE PERSPECTIVE
NEUROÉCONOMIQUE
PORTÉE SUR LES
INTERACTIONS
SOCIALES.
CHRISTIAN SCHMIDT
ÉTAIT VENU LE 24
SEPTEMBRE 2010
PRÉSENTER DANS
LE SÉMINAIRE
AEGIS SON LIVRE
PRÉCÉDENT.

Ce livre (Schmidt & Livet, 2014) fait suite à un premier, déjà consacré à la neuroéconomie (Schmidt, 2010 ; 2011). Il repose sur une rencontre entre un économiste et un philosophe et il a fallu trois années de dialogue entre les deux auteurs pour le produire. Le thème qu'il aborde soulève une question préalable : pourquoi adopter une perspective neuroéconomique pour comprendre les interactions sociales ? Deux objections peuvent être soulevées à l'encontre de ce projet. La première, qui émane plutôt des sociologues, consiste à soutenir que le social étant le social, on voit mal ce que pourrait apporter une science naturelle à la question des interactions. On peut répliquer qu'il existe aujourd'hui une sous-discipline des neurosciences, nommée précisément « neurosciences sociales », qui analyse, dans la perspective qui lui est propre, les comportements sociaux et contribue par conséquent à l'intelligence des interactions. Une seconde objection, un peu plus élaborée, reprend la vieille opposition entre le donné et l'acquis. De par leur enracinement dans la biologie, les neurosciences ont souvent tendance à rechercher dans la génétique l'origine des comportements

sociaux qu'elles analysent. C'est ainsi qu'elles s'efforcent, par exemple, de trouver des explications génétiques aux phénomènes de l'addiction. On observera toutefois à ce sujet que la génétique ne se contente plus aujourd'hui d'identifier les données, mais qu'elle s'intéresse à leur transformation dans une perspective dominée par l'évolutionnisme. Cela est d'autant plus vrai dans le cas du cerveau, dont l'étude a révélé les capacités de plasticité et de reproduction cellulaire. Non seulement, nous ne disposons pas de données une fois pour toutes, mais nous développons de nouvelles connexions neuronales, d'où la singularité du travail cérébral.

Il reste alors à montrer en quoi le recours aux neurosciences permet aujourd'hui de mieux



comprendre le fonctionnement des interactions sociales. Le livre dégage trois types de contributions de cette nouvelle neuroéconomie.

1. Le point de départ de ces recherches sur ces interactions est l'observation des comportements, le plus souvent par l'intermédiaire d'expériences. Cette observation révèle de nombreuses différences par rapport aux modèles de comportements supposés rationnels selon l'acception de l'économie traditionnelle. Pour expliquer ces comportements différents, il faut remonter aux conditions du fonctionnement cérébral qui les déterminent. C'est pourquoi la mise en relation de ces expériences avec ce que l'on peut connaître du fonctionnement du cerveau, grâce notamment aux possibilités offertes aujourd'hui par les techniques de l'imagerie cérébrale, fournit des informations essentielles qu'il nous faut alors interpréter.
2. Cette confrontation a conduit à reconsidérer plusieurs des découpages traditionnels de l'analyse des interactions. Il en va ainsi de la distinction entre « moi », « l'autre », « les autres ». On peut en effet identifier plusieurs types de « moi » et plusieurs types d'« autres », en fonction des positions respectivement occupées par le sujet et les autres sujets dans leurs interactions et selon la nature de ces interactions. Un certain nombre d'indicateurs pointent vers de nouvelles catégories.
3. De nouvelles distinctions catégorielles ont ainsi été mises en évidence. Elles portent les marques des phénomènes qui caractérisent ces interactions. Il s'agit, par exemple, de l'inter-intentionnalité qui anticipe la réciprocité et la conditionne. De même la confiance se décompose en deux sous-catégories, « la confiance/pari » et la « confiance/cadre ». Les normes implicites, qui prennent souvent la forme de conventions, doivent être distinguées des normes explicites. Quant aux normes explicites, elles doivent, pour pouvoir être mises en œuvre, se trouver complétées par des règles « mode d'emploi », elles-mêmes partiellement explicites.

Le livre est composé de trois parties : l'intersubjectivité traitée par Pierre Livet ; la coordination et la coopération analysées par Christian Schmidt et les normes et les règles développées par les deux auteurs. Chaque fois, l'un de nous deux a rédigé un chapitre et l'autre a réagi par son commentaire, parfois aussi développé que le chapitre lui-même.

Pour illustrer les apports du livre, quelques exemples peuvent être proposés. Ces développements ne consistent pas seulement à rendre compte d'expériences de neuroéconomie, ils visent à en tirer de nouvelles interprétations pour comprendre les interactions.

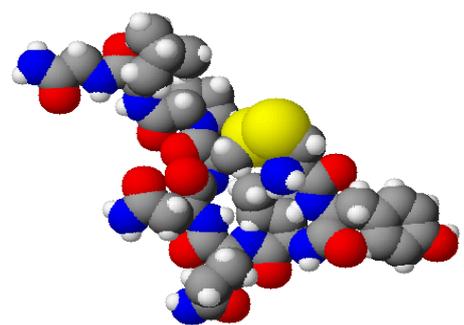
Considérons, par exemple, le *discount* temporel. Si j'ai le choix entre un gain attendu dans deux jours, et un gain attendu supérieur dans trois jours, je vais sans doute choisir le second. Mais si le futur est plus lointain, je vais le plus souvent préférer le gain proche. Cela doit évidemment être rapproché de l'incertitude qui accompagne la perception du futur. Mais l'écart entre les deux situations est plus fort que ne le laisserait supposer ce simple rapprochement. Pour comprendre cet écart nous suggérons de transposer ici un modèle moléculaire dans lequel on distingue les interactions à courte portée entre molécules et les interactions à longue portée qui correspondent à des processus d'interaction différents.

Autre manifestation du *discount* temporel. Je dois être en forme pour passer un concours. Il y a une soirée sympa de prévue. Je décide d'y aller mais d'en partir tôt. Au moment où je dois partir, je choisis de rester. Ce n'est pas seulement un plus accordé au présent : c'est que j'anticipe différentes interactions à moyenne portée dans la suite de la soirée, et même des interactions de second ordre entre ces interactions à moyenne portée. Le *trade-off* se fait entre une interaction à longue portée et ces interactions de second ordre. Il n'y a pas une solution unique au problème de savoir comment ordonner ces interactions. Le moyen de s'en sortir consiste alors à anticiper le regret que j'aurai quand j'arriverai au jour du concours sur les rotules. Mais cela requiert un effort mental considérable. Kahneman distingue, dans son dernier ouvrage, un système frugal rapide mais fruste et un système plus approfondi beaucoup plus lent, mais on voit que ce serait une simplification erronée de relier système fruste et interactions à courte portée et système approfondi et interactions à longue portée.

La distinction confiance cadre/confiance pari, tout d'abord. Les résultats de neuroéconomie avec injections d'ocytocine montrent que cela marche mieux quand le sujet produit lui-même l'ocytocine (qui se produit dans la relation entre le nourrisson et sa mère lors de l'allaitement), que quand on la lui injecte, cas dans lesquels il peut y avoir agressivité. En réalité, la confiance semble produire de l'ocytocine, plus que l'ocytocine ne produit de la confiance.

Ensuite, le plaisir de punir (expérience de Quervain *et al.*, 2004). On est dans une relation d'agence, avec un investissement. Un mandant confie une somme à un mandataire, et celui-ci lui rend la somme avec un surcroît. Si le surcroît est trop faible, le mandant peut punir. En majorité, les personnes de l'expérience punissent. La conclusion des auteurs de l'expérience est que, puisque la punition est coûteuse pour les deux, il existe un plaisir à punir. Une autre interprétation est possible : la reconnaissance sociale des torts. Les tiers ont reconnu le dommage puisqu'ils me permettent de punir, et cette reconnaissance active le système de la récompense.

Enfin, le rôle des règles explicites. À quoi servent-elles ? On peut repartir de l'aporie de Wittgenstein : comment peut-on savoir qu'un comportement passé qui semble avoir suivi une règle assure que ce sera bien cette règle qui sera suivie dans le futur ? On ne peut recourir à des règles explicites pour le résoudre, puisque s'il faut une règle pour appliquer une règle, la régression sera infinie. On utilise donc des modes d'emploi, c'est-à-dire des usages qui donnent quelques repères pour des règles. Les modes d'emploi, en sélectionnant des traits repérables des usages, permettent des contrôles (là, on est plus proche de Foucault que de Wittgenstein). Les sociétés mettent en place différents niveaux de contrôle. Un premier contrôle porte sur la fiabilité de la communication, qui est nécessaire aux coordinations, faites dans l'intérêt de chacun. Elles diffèrent des coopérations, qu'il peut être dans l'intérêt de certains de tenter d'exploiter. Or même une fois qu'une coopération est en route, il existe plusieurs manières de se coordonner pour coopérer, et il faut donc pouvoir disposer d'un deuxième mode de contrôle, qui vérifie cette coordination des manières de coopérer. Un mode de contrôle encore plus poussé consiste à se donner des repères de la conformité des conduites avant résultats. Nous sommes très sensibles à ces indices donnés par nos manières d'orienter nos conduites au cours de nos engagements et nous tentons de minimiser ces engagements, comme le montrent des analyses des diverses conditions expérimentales étudiées pour la situation dite du



La molécule d'ocytocine

« trolley fou », entre lesquelles nous préférons celles qui minimisent ces interventions en cours d'engagement, et donc notre responsabilité. Pour couronner cet édifice social des modes de contrôle, les institutions juridiques et politiques ont même mis en place un contrôle des capacités des règles explicites à assurer ces contrôles.

Revenons sur la question posée par les observations expérimentales concernant le traitement du *discount* temporel. Si on le considère dans la perspective de l'analyse économique, il s'agit plutôt d'approfondir la nature des relations entre l'appréhension par les sujets de la distance temporelle et du risque, relations confirmées par les informations dont on dispose sur le travail cérébral dans les deux cas. L'hypothèse sur laquelle nous travaillons aujourd'hui est celle d'une structure fondamentale de la perception temporelle qui se traduirait selon deux modalités différentes correspondant respectivement à la distance temporelle par rapport au point de référence et au risque.

Passons à la notion d'inter-intentionnalité. L'objet de la théorie des jeux est de fournir une modélisation logique et claire des interactions. Malheureusement, beaucoup d'expériences montrent que dans plusieurs situations d'interactions les joueurs ne se comportent pas comme le voudrait la théorie des jeux. C'est la raison pour laquelle nous proposons de rechercher les explications de cet écart au niveau du fonctionnement cérébral dans ces situations d'interactions. En voici un exemple bien connu. Dans le jeu de l'*ultimatum*, l'un des joueurs reçoit une somme d'argent qu'il lui revient de partager à sa convenance avec un autre joueur qui peut accepter ou refuser ce partage. Si ce dernier accepte, la somme sera partagée selon les conditions fixées par le premier joueur. Dans le cas contraire cette somme sera retirée du jeu et aucun des deux joueurs n'en bénéficiera. En théorie des jeux, le premier joueur a tout intérêt à proposer un partage où il garderait pour lui l'essentiel de la somme car le second joueur a toujours intérêt à accepter le partage. Même faible, ce gain sera supérieur au résultat de son refus. C'est ce qui fonde ici l'équilibre de Nash qui avantage le premier joueur. Des anthropologues ont testé le jeu dans différentes cultures et, sauf dans une tribu reculée du Pérou, personne ne s'est jamais comporté comme le voudrait la stratégie Nashienne (Henrich, 2000). On a également organisé des jeux dans lesquels le joueur en premier jouait contre un ordinateur doté du programme de Nash. Les résultats obtenus et, plus intéressant encore pour nous, les activations cérébrales se sont révélées différents selon que le joueur jouait face à un autre joueur ou face à un ordinateur. En amont du jeu stratégique ancré à un niveau cérébral plus profond, il existerait donc une perception sociale. Pour la théorie des jeux, cela veut dire que nous raisonnons toujours avec des attentes de l'autre et sous l'effet de l'émotion qu'elles suscitent en nous. Nous sommes certes rationnels, mais notre cerveau fonctionne dans un système plus riche fait de sensations, d'impressions et d'émotions. Les données traitées par la théorie des jeux sont donc insuffisantes pour comprendre les véritables mécanismes des interactions sociales. Les théoriciens ont eux-mêmes perçu cette insuffisance à laquelle ils ont tenté de trouver des solutions. Ils se sont appuyés sur l'hypothèse d'un *common knowledge* ou ont inventé des jeux psychologiques (Geanakoplos *et al.*, 1989 ; Rabin, 2013). Aumann & Brandenburger (1995) avaient, de leur côté, exploré les problèmes de croyances réciproques. Mais ces croyances se construisent et évoluent en dynamique, ce qui veut dire que le modèle dynamique pour en rendre compte n'est pas une simple extension dans le temps du jeu statique.

Le problème principal posé par nos interactions réside en définitive dans la représentation de l'autre comme semblable et différent. Comment puis-je me mettre

à la place de l'autre joueur (d'un autre joueur) ? Petit à petit, il faut pouvoir élaborer une perception de l'autre comme différent au cours du déroulement du jeu, c'est-à-dire grâce au caractère interactif de la situation de jeu. L'hypothèse de l'inter-intentionnalité part de la découverte que notre système cérébral formule des hypothèses sur les intentions de l'autre dès le déclenchement de toute interaction. Les cas d'asymétrie de positions (l'investisseur qui confie une somme à un mandataire) mettent ainsi en évidence les différences entre une intentionnalité prêtée *a priori* à l'autre et l'intentionnalité de l'autre induite de l'observation de son comportement. Les perceptions de risque ne sont pas les mêmes dans les deux cas, d'où la différence entre le risque-pari et le risque-cadre précédemment évoquée. C'est à partir de l'inter-intentionnalité que l'on pourra introduire la notion de réciprocation. Elle nous permet de faire une distinction plus claire entre la coordination et la coopération. Comment passe-t-on de l'interaction à la coopération ? Les réponses traditionnelles ne sont pas satisfaisantes et la réciprocation en fournit une explication plus convaincante. L'inter-intentionnalité s'exprime par la réciprocation. On peut, en effet, faire crédit au départ d'une inter-intentionnalité, mais c'est le plus souvent dans l'attente d'une réciprocation. Si cette dernière se manifeste, alors la coopération s'instaure. Certains théoriciens des jeux acquis à cette formulation ont même essayé de définir en ces termes un équilibre de réciprocation (Falk & Fischbacher, 2006). La coopération se fait alors de manière conditionnelle, à cette condition.

DÉBAT

Anni Borzeix : en vous écoutant parler de la coopération conditionnelle, on pense au système de don et de contre-don de Mauss. N'y a-t-il pas un pont à construire avec l'anthropologie ?

Réponse : Nous avons regardé ce point. L'interaction se joue dans les deux sens. Il existe plusieurs interprétations possibles du système en termes de théorie des jeux. Marcel Hénaff (2012) a essayé de donner des réinterprétations plus fines, du côté de l'anthropologie. Et on a donc une interaction (pour le coup) entre anthropologie et théorie des jeux.

Il y a eu beaucoup de travaux sur les écarts entre résultats de jeux selon les sociétés, des différences ont été perçues. Les neurosciences ont tendance à dire que tous les cerveaux fonctionnent de la même manière. Mais la nuance vient du fait que les cerveaux ont été formés de manière un peu différente et du coup fonctionnent un peu différemment. C'est possible dans une perspective évolutionniste.

Dans les jeux évolutionnistes, on voit que le principe de maximisation de l'utilité ne fonctionne pas, c'est plutôt la minimisation du regret.

Hervé Dumez : Pouvez-vous revenir sur les apports en termes d'analyse de la confiance ?

Réponse : L'investisseur, s'il envoie de l'argent, manifeste une certaine confiance. Mais elle n'est pas déterminée par un cadre fixé par des interactions précédentes. C'est la confiance-pari. Si nous entrons dans un milieu social que nous connaissons mal, nous avons cette confiance-pari. Les interactions qui se produisent permettent d'entrer dans une confiance-cadre. Il est intéressant de s'interroger sur la manière

dont on passe de l'une à l'autre. Le cadre est une construction au cours du jeu. On pourrait imaginer un degré de pari qui serait une rupture sur le cadre.

Hervé Dumez : Dans l'histoire des échecs, on raconte qu'un maître qui jouait contre plusieurs adversaires en même temps a été confronté avec lui-même lors d'un tournoi (alors qu'il avait les noirs dans une partie et les blancs dans une autre) ; a-t-on fait des expériences de confrontation d'un cerveau face à lui-même ?

Réponse : Cela rejoint l'idée de l'autre moi-même, qui se trouve déjà dans Schelling. Il y a deux effets. Les expériences montrent que l'on ne construit son identité que dans l'interaction et quand il y a un retour face à soi-même. Il existe une différence essentielle entre l'interaction avec les autres et la construction de l'identité. J'ai une dynamique avec mon propre passé, qui est alimentée par des interactions avec des éléments de mon environnement. Et avec les autres, deux dynamiques se co-construisent. On peut en donner une preuve indirecte. Vous êtes en interaction avec un joueur, ou vous suivez les interactions de deux joueurs, et les deux jeux sont les mêmes. Cependant le fait de participer au jeu est très différent de la simple observation.

Alain Jeunemaitre : Plusieurs questions. D'abord pouvez-vous revenir sur la construction des séquences des jeux. Autre point : concernant l'inter-intentionnalité, par exemple dans le cas la vente des produits financiers sophistiqués, que se passe-t-il quand il n'y a pas d'équilibre compréhensif entre le vendeur et l'acheteur ? À propos de la coopération, ne peut-il pas y avoir des phénomènes de réciprocité très différents ? Dernière question : que peut-on dire de l'intelligence artificielle ? On peut aujourd'hui faire dialoguer des robots et développer de l'apprentissage entre robots.

Réponse : (À propos de la *behavioral finance*, Christian Schmidt revient à la critique) les comportements ne correspondent pas aux modèles habituels de la rationalité. Ce que les neurosciences ont bien mis en évidence, c'est le fait que l'ambiguïté est un phénomène en soi. Soit on n'a pas assez d'information, soit on a des informations contradictoires. Ce qui est intéressant dans le cas des marchés financiers est que l'ambiguïté est la règle, non l'exception. Dans ces situations-là, le cerveau s'active de manière très particulière : il n'applique pas un modèle, il recherche, par exemple dans la mémoire, parfois dans une mémoire de l'imaginaire. La première réaction du cerveau face à l'ambiguïté est l'aversion. On est plus aversé à l'ambiguïté qu'au risque. Du

coup, dans les situations d'ambiguïté, le cerveau va prendre des risques, parce qu'il préfère se trouver dans une situation de risque. La contradiction est qu'en situation d'ambiguïté, il faudrait réfléchir. Mais on n'a pas ce temps sur les marchés financiers, donc on prend des risques. J'ai écrit un papier sur le sujet dans la revue des analystes financiers « Des finances comportementales aux neurofinances ». Du point de vue application, la finance est le domaine qui se dégage fortement. La finance est en désarroi et s'intéresse du coup beaucoup aux neurosciences.

Il faut se méfier des métaphores, le cerveau ne fonctionne pas du tout comme un ordinateur. D'où un problème supplémentaire : la relation



La règle de réciprocation
<http://bizztro.tumblr.com/post/77681000441/the-psychology-of-reciprocation>

entre le cerveau et les machines. La question sur les robots renvoie à la question des réseaux sociaux. Nous sommes conscients de n'avoir pas poussé assez loin la question du fonctionnement de la coopération selon la structure des réseaux. La structure des réseaux a sans doute un effet sur la coopération.

Il existe des expériences dans lesquelles on met un sujet en position d'observateur, puis on le met en situation d'acteur. Les configurations cérébrales sont alors très différentes. Et les rôles du tiers peuvent varier (il analyse simplement ou il a la capacité de punir les joueurs). Certaines coopérations ne se déclenchent pas s'il n'y a pas un tiers. Par ailleurs, dans les expériences, les joueurs peuvent avoir des anticipations sur ce qu'attend l'expérimentateur.

Il y a deux altérités, celle de l'autre avec lequel on joue, et il y a un grand Autre (si l'on veut paraphraser Lacan), un autre en surplomb, qui peut en fait se décliner de manières variées selon les différents rôles que peuvent jouer les tiers ■

Références

- Aumann Robert J. & Brandenburger Adam (1995) "Epistemic conditions for Nash equilibrium", *Econometrica*, vol. 63, n° 5, pp. 1161-1180.
- Falk Armin & Fischbacher Urs (2006) "A theory of reciprocity", *Games and Economic Behavior*, vol. 54, n° 2, pp. 293-315.
- Geanakoplos John, Pearce David & Stacchetti Ennio (1989) "Psychological games and sequential rationality", *Games and Economic Behavior*, vol. 1, n° 1, pp. 60-79.
- Hénaff Marcel (2012) *Le don des philosophes. Repenser la réciprocité*, Paris, Le Seuil.
- Henrich Joseph (2000) "Does culture matter in economics behavior? Ultimatum bargaining among Machiguenga of the Peruvian Amazon", *The American Economic Review*, vol. 90, n° 4, pp. 973-979
- Quervain Dominique J.F. de, Fischbacher Urs, Treyer Valérie, Schellhammer Melanie, Schnyder Ulrich, Buck Alfred & Fehr Ernst (2004) "The neural basis of altruist punishment", *Science*, vol. 305, n° 5688, pp. 1254-1258.
- Rabin Matthews (2013) "An approach to incorporating psychology into economics" *The American Economic Review*, vol. 103, n° 3, pp. 617-622.
- Schmidt Christian (2010) *Neuroéconomie : comment les neurosciences transforment l'analyse économique*, Paris, Odile Jacob.
- Schmidt Christian (2011) "La neuroéconomie en question", *Le Libellio d'Aegis*, vol. 7, n° 1, pp. 15-27.
- Schmidt Christian (2015) "De la finance comportementale à la neurofinance", *Analyse financière*, n°54, pp. 81-84.
- Schmidt Christian & Livet Pierre (2014) *Comprendre nos interactions sociales. Une perspective neuroéconomique*, Paris, Odile Jacob.