

DU ROLE DES EMOTIONS ET DES LIENS SOCIAUX

DANS LES JEUX DE BIEN PUBLIC : ETUDES

COMPORTEMENTALES ET NEUROECONOMIQUES

Benjamin Pelloux

Résumé étendu en français

Cette thèse présente quatre essais sur le rôle des émotions et leur influence sur la prise de décision ainsi que sur le développement dynamique de liens affectifs entre des parties. Plus précisément, les trois premiers chapitres s'intéressent à la création et l'évolution de liens sociaux au cours d'interactions anonymes en laboratoire. Le problème est abordé aux niveaux théorique, comportemental et neural de telle sorte que nous proposons une vue d'ensemble assez complète sur la façon dont les liens sociaux sont créés et dont ils guident les décisions lors d'interactions économiques répétées. Le dernier chapitre explore l'utilisation de *feedback* verbal en tant qu'outil visant à augmenter la coopération entre individus ainsi que le rôle sous-jacent joué par les émotions dans cet environnement.

Les liens sociaux sont des liens affectifs se créant entre des agents en interaction, les faisant se préoccuper du sort de chacun et donc influençant leurs actions. Ces liens surviennent uniquement à travers l'occurrence d'une interaction sociale et sont influencés par les sentiments provoqués par l'interaction. Par conséquent, ces liens sociaux peuvent être positifs ou négatifs en fonction des émotions ressenties et de la satisfaction dérivée du comportement de l'autre partie. Il a été montré que ce « besoin d'appartenance » (Baumeister et Leary, 1995) joue un rôle prépondérant dans la vie humaine et apparaît comme l'une des

principales motivations sous-jacentes à notre comportement social. Des situations économiquement pertinentes aussi diverses que le travail d'équipe, la vie de quartier, les oligopoles ou les relations acheteur-vendeur sont toutes susceptibles de créer des liens entre les parties impliquées. Ce dernier point peut être illustré par une citation tirée de la chanson « Affreux, bête et dangereux » du rappeur français Ekoué: « Les sociologues, les politologues qui viennent chez nous pour acheter de la drogue appellent ça créer du lien social ». Il s'agit en effet d'un exemple illustratif de la définition des liens sociaux : l'occurrence d'une interaction (ici, un échange économique) donne lieu à un contact entre des personnes jusque-là anonymes l'une pour l'autre et pourrait créer un lien social entre les parties interagissant. Découlant également de cette définition, la qualité de l'attachement (positif ou négatif) de chaque partie dépendra de la satisfaction retirée de l'interaction qui ne saurait être toujours symétrique, comme suggéré implicitement par la citation.

Les dilemmes sociaux sont des situations dans lesquelles il est dans l'intérêt du groupe de coopérer et de partager les bénéfices de la coopération mais également où chaque individu a en même temps une incitation à profiter de ces bénéfices sans subir le coût de la coopération. La trame commune de cette thèse est le conflit très simple entre incitations individuelles et sociales impliqués par les dilemmes sociaux. Les jeux de bien public, une forme généralisée du dilemme du prisonnier, offrent un tel cadre d'analyse. Aussi simple que ce jeu peut paraître, il a déclenché un champ de recherche très riche et presque infini tant les réponses apportées semblent provoquer de nouvelles questions. De nombreux résultats expérimentaux ont été réunis par les chercheurs au cours d'expériences en laboratoire et plusieurs régularités empiriques ont émergé (Ostrom, 1995; Ledyard, 1995; Chaudhuri, 2011). Plus important, il apparaît que, contrairement aux prédictions de la théorie des jeux standard¹, les individus contribuent initialement une part significative de leur richesse au bien public (même lorsqu'ils

¹ Au cours de cette thèse, nous utiliserons l'expression « Théorie des jeux standard » en faisant référence à la théorie des jeux qui suppose des agents purement égoïstes et rationnels contraints par aucune limite cognitive.

sont rappareillés anonymement entre chaque répétition du jeu) avant un déclin progressif des contributions. Plusieurs outils ont été utilisés afin d'augmenter le niveau de coopération comme la punition (Fehr and Gächter, 2000) ou la récompense couteux (Sefton and Steinberg, 1996), les sanction symboliques (Masclot et al., 2003), la communication avant le jeu (Bochet et al., 2006), l'implémentation de leadership au sein des groupes (Güth et al., 2007) ou la sélection endogène des autres membres du groupe (Page et al., 2005).

Dans le but de résoudre les problèmes de la théorie des jeux standards à prédire la coopération des individus, même lors d'interactions non répétées, les chercheurs ont développé des théories incluant différents types de préférences regroupées sous le terme général de « préférences sociales ». Ces préférences sont modélisées sous la forme de fonctions d'utilité interdépendantes, ce qui signifie qu'un individu valorise, de manière positive ou négative et ce en fonction de différents facteurs, le bien-être des autres personnes avec qui il interagit (Sobel, 2005). Des modèles ont été développés afin de refléter des motivations guidant ces préférences sociales comme par exemple, la réciprocité (Rabin, 1993), l'aversion à l'inégalité (Fehr and Schmidt, 1999 ; Bolton et Ockenfels, 2000), l'efficacité (Charness et Rabin, 2002) ou l'aversion à la culpabilité (Battigalli et Dufwenberg, 2007). Ces modèles supposent tous que ces préférences sociales sont basées sur des traits de personnalité stables et abordent de façon très cognitive et basée sur les croyances le comportement humain dans les dilemmes sociaux.

Une autre façon de modéliser la manière dont nous interagissons avec d'autres personnes est d'utiliser les fonctions d'utilité interdépendantes afin de formaliser le développement de liens affectifs entre les différentes parties. Comment en venons-nous à nous soucier des étrangers avec qui nous interagissons ? Un mécanisme très simple serait de commencer à accorder un poids au bien-être d'autrui en fonction de la satisfaction retirée de l'interaction que nous avons eu avec eux. Une première formalisation du développement des

liens sociaux dans un environnement de bien public a été développé par van Dijk et van Winden (1997). L'idée sous-jacente est que notre attitude envers les autres se construit au fur et à mesure des occurrences de l'interaction, étant donc intrinsèquement dynamique. Cette attitude peut-être positive (ex. amitiés) ou négative (ex. relation difficiles au sein de mauvais voisinages), n'est pas nécessairement symétrique (ex. une passion non réciproque) et évolue sur la base des sentiments ressentis au cours de l'interaction. En dernier lieu, notre sentiment envers l'autre guidera nos actions lorsque nos décisions auront des conséquences sur la personne considérée.

Nous utilisons le mot sentiment ici car la littérature assigne un rôle important aux émotions, en faisant l'une des principales forces guidant l'attachement social. La formalisation de l'agent économique en tant qu'*homo economicus* attribué au lui attribue des capacités de calcul et de raisonnement infinies, lui permettant de toujours choisir ce qui est le mieux pour lui en fonction de ces préférences dans n'importe quel environnement ou situation. La domination d'une telle théorie définissant les agents comme des décisionnaires froids et rationnels peut expliquer pourquoi les émotions ont été absentes de la plupart des théories du comportement économique jusqu'à récemment. Cependant, de plus en plus de recherches sont désormais effectuées dans le but de découvrir les déterminants économiques du comportement. Si cette ligne de recherche a commencé par se concentrer sur des questions liées à la prise de décision individuelle au travers du développement de théories telles que l'aversion au regret (Loomes et Sudgen, 1982) ou l'aversion à la perte (Tversky et Kahneman, 1992), les chercheurs ont maintenant commencé à incorporer les émotions dans leurs théories sur la prise de décision sociale (ex. aversion à la culpabilité, aversion à la trahison, aversion au mensonge, etc...). De manière plus empirique, le rôle joué par les émotions dans les dilemmes sociaux a été étudié, particulièrement dans des jeux permettant la punition et les représailles (voir par exemple Hopfenstiz et Reuben, 2009). Notre travail considère le rôle des

émotions à la fois de manière empirique mais également au travers du développement d'un modèle théorique de développement dynamiques des liens sociaux.

Comme le suggère son titre, cette thèse emploie différents outils méthodologiques afin d'étudier le rôle des émotions et des liens sociaux dans les jeux de bien public. Des procédures expérimentales comportementales sont utilisées tout le long de cette thèse et toutes les données présentées ici furent collectées à l'intérieur du laboratoire. Ceci nous permet de contrôler au maximum l'environnement où évoluent les agents et de contrôler pour le degré d'anonymat les séparant. Dans les chapitres 2 à 4, nous considérons spécialement les liens sociaux tissés en laboratoire entre deux anonymes et ne sommes pas intéressés par des relations préexistantes. Souhaitant aussi étudier l'impact des émotions sur les décisions, nous utilisons des auto-évaluations afin de les mesurer. Même si elles sont rarement utilisées en sciences économiques, les auto-évaluations des émotions apparaissent comme un outil de valeur pour les mesurer en comparaison à des techniques plus invasives. Plusieurs chercheurs ont montré qu'elles étaient fiables et, pour certaines émotions sociales complexes comme la honte ou la culpabilité, inévitables (Adolphs, 2002; Tangney and Dearing, 2002; Ortony et al., 1988).

Si les expériences comportementales en laboratoire sont désormais largement acceptées en tant que méthode utilisable afin de récolter des données sur le comportement économique, l'utilisation d'outils neuroscientifiques afin d'étudier les fondements neuraux de la prise de décision est encore loin d'être reconnue comme précieuse par les économistes. Le développement d'études économiques utilisant l'Imagerie à Résonance Magnétique fonctionnelle (IRMf) a provoqué un intense débat au sein des économistes durant lequel des positions assez extrêmes ont été prises quant à la définition de ce qu'est la neuroéconomie et

de ce qu'elle pourrait apporter dans le futur², à la fois parmi les critiques et les louanges. Par exemple, Gul et Pesendorfer (2008) rejettent toute utilisation potentielle des outils neuroscientifiques en sciences économiques « car ces dernières ne font aucune hypothèse et ne tirent aucune conclusion quant à la physiologie et le cerveau » (op. cit., p4) et semblent discréditer tout travail pluridisciplinaire entre économistes et psychologues puisque « les sciences économiques et la psychologie s'intéressent à des questions différentes, utilisent des abstractions différentes et utilisent différents types de preuves empiriques (op. cit., p4). De manière plus pratique, d'autres ont critiqué les méthodes et les statistiques de la neuroéconomie sans en renier le potentiel futur (Harrison, 2008). De l'autre côté du spectre, Camerer et al. (2005) avance que « les avancées en neurosciences rendent désormais possible une mesure directe des pensées et des sentiments, ouvrant ainsi la 'boîte noire' qui est la pierre angulaire de n'importe quel système ou interaction économique : l'esprit humain » (op. cit., p53). Le contenu de cette thèse tout comme le parcours des auteurs impliqués dans certains des chapitres indique que nous valorisons le travail pluridisciplinaire et que nous croyons qu'économistes et psychologues peuvent apprendre beaucoup les uns des autres, au niveau conceptuel comme au niveau méthodologique. Cependant, nous nous éloignons de la vue selon laquelle les neurosciences permettraient de lire « les pensées et les sentiments ».

D'abord, nous voyons la neuroéconomie comme une source additionnelle de preuves concernant la prise de décision économique et ces processus motivationnels sous-jacents. Certains modèles économiques de choix formulent des hypothèses sur ces derniers et nous pensons que la neuroéconomie permet de ressembler des résultats permettant d'instruire ces modèles. Ensuite, le fait que la neuroéconomie puisse être utile ne signifie pas que les expériences comportementales doivent être abandonnées. Il faut avoir une idée bien précise quant au comportement des sujets de désirer étudier les mécanismes neuraux supportant ces

² Pour une vue d'ensemble de ce débat, voir par exemple Rustichini (2005), Bonanno et al. (2008), Fehr et Rangel (2011) et Pelloux et al. (2009).

décisions.³ Il nous apparaît nécessaire, afin de générer des hypothèses pertinentes sur les processus neuraux et psychologiques sous-tendant la prise de décision, de raffiner au maximum nos connaissances sur les causes, les motivations, les processus mentaux, les caractéristiques environnementales et les différences individuelles qui engendrent le comportement observés. L'absence d'un tel travail préliminaire pourrait mener à une mauvaise interprétation des activations neurales et conduire à des conclusions erronées. Finalement, nous reconnaissons que les chercheurs doivent rester prudents lors de l'interprétation de leurs résultats, tout spécialement quand ils concernent des processus neuraux de niveau élevé comme ceux qui sont requis pour prendre des décisions économiques. La différence entre ce qui constitue les données et les extrapolations et conjectures qui sont faites à partir de ces données doit être rendue claire aux lecteurs.

Le chapitre 2 commence par étendre l'analyse théorique des liens sociaux développée par van Dijk et van Winden (1997). Nous étendons ce modèle en incluant un mécanisme de choix probabiliste et, surtout, en permettant aux agents de prendre en compte le futur. L'objectif principal de ce chapitre est de confronter ce modèle théorique de choix aux données réunies en laboratoire. Nous utiliserons dans ce but des données issues de plusieurs expériences de jeux de bien public qui diffèrent en termes de structure du jeu et du nombre de parties impliquées. Notre objectif empirique est double dans ce chapitre. D'abord, nous estimons le modèle de lien social au niveau du groupe et de l'individu et montrons qu'il a un pouvoir explicatif certain sur les données et que des liens sociaux sont effectivement développés. Ces résultats suggèrent que l'histoire de l'interaction joue un rôle prépondérant dans la détermination des préférences, que de la réciprocité a lieu et qu'environ la moitié des sujets sont prospectifs. Deuxièmement, nous générons des prédictions comportementales

³ C'est bien ainsi que nous avons procédé puisque de nombreuses expériences liées au développement des liens sociaux dans des jeux économiques ont été effectuées avant notre étude de ces problèmes au cours des chapitres 3 et 4.

individuelles afin de suivre les changements de contributions de périodes en périodes. La performance du modèle apparaît bonne, celui-ci étant capable de prédire schémas dynamiques complexes. Ceci étant vrai pour les jeux joués en paires comme au sein de groupes de quatre joueurs où les processus de prise de décision sont bien plus complexes.

Suivant cette approche comportementale et dans le but de rassembler différents types de preuves, les chapitres 3 et 4 s'intéressent aux mécanismes cérébraux causant le comportement lors de ces expériences. Dans le chapitre 3, nous examinons comment les différences individuelles en termes d'empathie et de lien social modulent les réponses neurales à un partage imposé de ressources monétaires. Premièrement, nous trouvons que ce partage entraîne l'activation de systèmes neuraux associés à la récompense (striatum) et l'empathie (cortex insulaire antérieur et cortex cingulaire antérieur) seulement après l'occurrence d'une interaction, et non avant. Nous étendons ainsi de précédents résultats (Singer et al. 2004 ; 2006) et montrons que les réponses neurales liées à l'empathie ne surviennent qu'après une interaction même si celle-ci est aussi peu intense qu'une interaction anonyme de type économique peut l'être. Deuxièmement, le partage provoque aussi l'activation de zones associées à l'altruisme et la significativité sociale (sulcus temporal supérieur postérieur, pSTS) corrélant avec le succès de l'interaction et l'appréciation du partenaire après l'interaction. Finalement, nous montrons que seulement l'activité cérébrale liée au lien social prédit le comportement pro-social lors des périodes suivantes de l'interaction, suggérant le pSTS comme substrat neural du lien social.

Dans le chapitre 4, nous utilisons l'IRMf afin de trouver les corrélats neuraux des éléments du modèle théorique présenté lors du chapitre 2 et de confirmer le rôle du pSTS suggéré par le chapitre 3. Les résultats montrent qu'au moment de la prise de décision, l'activité dans le pSTS et la jonction temporo-pariétale (TPJ) corrèle avec l'estimation du lien social, supportant une fois de plus le rôle joué par ces régions dans la formation du lien social.

De plus, l'activité dans le cortex préfrontal médian (mPFC, associé à des fonctions cognitives plus élevées comme la planification et la prise de décision à long terme) corrèle avec la contribution au bien public. En outre, nous pouvons supporter le rôle des émotions dans le mécanisme de formation du lien social : lorsque l'information sur la période précédente est transmise, le degré de coopération de l'autre joueur est encodé dans des régions précédemment impliquées dans les émotions liées aux récompenses, tout spécialement dans des contextes sociaux (striatum, cortex insulaire antérieur, cortex cingulaire antérieur, pSTS, TPJ). Finalement, nous trouvons une connectivité fonctionnelle entre le pSTS et le mPFC, suggérant le fait que la représentation du lien social soit intégrée dans le processus de décision.

Enfin, le chapitre 5 étudie l'usage potentiel de la communication verbale comme outil de *feedback* afin de soutenir la coopération dans les jeux de bien public. Du fait de sa haute charge émotionnelle et de son faible coût en comparaison à la punition monétaire, nous avons considéré le feedback verbal comme un mécanisme possiblement utile afin de promouvoir le comportement pro-social comme cela est suggéré par d'autres résultats dans des jeux différents (Xiao et Houser, 2005 ; 2009 ; Ellingsen et Johannesson, 2008). Nous trouvons que ce mécanisme est d'une efficacité limitée comparé à la sanction monétaire. Il apparaît capable de maintenir un niveau de coopération plus élevé (comparé au traitement de contrôle sans communication ni sanction) pour un moment mais pas de le maintenir sur le long terme. Il est intéressant de noter que provoquer de la culpabilité encourage la coopération mais que cet effet disparaît sur le long terme. Au contraire, la provocation de réactions colériques a un impact négatif sur la coopération.

Cette thèse combine plusieurs méthodologies afin d'examiner les déterminants du comportement lors de jeux de bien public répétés. Nous avons montré que les préférences sociales dynamiques modélisés par le développement de liens affectifs entre différentes

parties étaient performantes afin d'expliquer le comportement de contribution dans les jeux de bien public. De plus, une utilisation ambitieuse de l'IRMf alliée à un modèle théorique de prise de décision nous a permis de trouver les substrats neuraux des paramètres de notre modèle ainsi que des corrélats neuraux du lien social cohérents entre différentes tâches. Ces résultats indiquent la présence d'un réseau impliqué dans la prise de décision dans un environnement de bien public. Cette thèse éclaire aussi d'une nouvelle lumière l'impact de la communication sur le comportement coopératif et ses processus émotionnels sous-jacents.

Références:

- Adolphs, R. (2002): "Recognizing Emotion From Facial Expressions: Psychological and Neurological Mechanisms," *Behavioral and Cognitive Neuroscience Reviews*, 1, 21-61.
- Battigalli, P., and M. Dufwenberg (2007): "Guilt in Games," *American Economic Review*, 97(2), 170-176.
- Baumeister, R. F., and M. R. Leary (1995): "The need to belong: Desire for interpersonal attachments as a fundamental human motivation," *Psychological Bulletin*, 117(3), 497-529.
- Bochet, O., T. Page, and L. Putterman (2006): "Communication and Punishment in Voluntary Contribution Experiments," *Journal of Economic Behavior & Organisation*, 60(1), 11-26.
- Bolton, G. E., and A. Ockenfels (2000): "A theory of equity, reciprocity and competition," *American Economic Review*, 90(1), 166-193
- Bonanno, B., C. List, B. Tungodden, and P. Vallyentyne (2008): "Introduction to the Special Issue of Economics and Philosophy on Neuroeconomics," *Economics and Philosophy*, 24, 301-302.
- Camerer, C. F., G. Loewenstein, and D. Prelec (2005): "Neuroeconomics: How Neuroscience Can Inform Economics," *Journal of Economic Literature*, 43(1), 9-64.
- Charness, G., and M. Rabin (2002): "Understanding Social Preferences with Simple Tests," *The Quarterly Journal of Economics*, 117 (3), 817-869.
- Chaudhuri, A. (2011): "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature," *Experimental Economics*, 14(1), 47-83.

- van Dijk, F., and F. van Winden (1997): "Dynamics of social ties and local public good provision," *Journal of Public Economics*, 64(3), 323-341.
- Ellingsen, T., and M. Johannesson (2008): "Anticipated verbal feedback induces altruistic behavior," *Evolution and Human Behavior*, 29, 100-105.
- Fehr, E., and S. Gächter (2000): "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, 90(4), 980-994.
- Fehr, E., and K. M. Schmidt (1999): "A Theory of Fairness, Competition, and Cooperation," *The Quarterly Journal of Economics*, 114(3), 817-868.
- Gul, F., and W. Pesendorfer (2008): "The Case for Mindless Economics," in *The Foundations of Positive and Normative Economics*, eds Caplin, A., and A. Schotter, (New York: Oxford University Press), 3–39.
- Güth, W., V. Levati, M. Sutter, and E. van der Heijden (2007): "Leading by example with and without exclusion power in voluntary contribution experiments," *Journal of Public Economics*, 91, 1023-1042.
- Harrison, G. W. (2008): "Neuroeconomics: A Critical Reconsideration," *Economics & Philosophy*, 24, 303–344.
- Hopfensitz, A., and E. Reuben (2009): "The Importance of Emotions for the Effectiveness of Social Punishment," *Economic Journal*, 119, 1534-1559.
- Ledyard, J.O. (1995): "Public Goods: A Survey of Experimental Research," in *The Handbook of Experimental Economics*, eds. A. Roth & J. Kagel. (Princeton, NJ: Princeton University Press), 111–181.

- Loomes, G., and R. Sudgen (1982): "Regret Theory: An Alternative Theory Of Rational Choice Under Uncertainty," *The Economic Journal*, 92(368), 805-824.
- Masclet, D., C. Noussair, S. Tucker, and M-C. Villeval (2003): "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism," *American Economic Review*, 93(1), 366-380.
- Ortony, A., G. L. Clore, and A. Collins (1988): *The Cognitive Structure of Emotions*, Cambridge University Press.
- Ostrom, E. (2000): "Collective Actions and the Evolution of Social Norms," *Journal of Economic Perspective*, 14(3), 137-158.
- Page, T., L. Putterman, and B. Unel (2005): "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry and Efficiency," *The Economic Journal*, 115, 1032–1053.
- Pelloux, B., J.-L. Rullière, and F. van Winden (2009): "La Neuroéconomie dans l'Agenda de l'Economie Comportementale," *Revue Française d'Economie*, 23(4), 3-36.
- Rabin, M. (1993): "Incorporating Fairness into Game Theory and Economics," *American Economic Review*, 83(5), 1281-1302.
- Rustichini, A. (2005): "Neuroeconomics: Present and future," *Games and Economic Behavior*, 52, 201-212.
- Sefton, M., and R. Steinberg (1996): "Reward structures in public good experiments," *Journal of Public Economics*, 61(2), 263-287.

- Singer, T., B. Seymour, J. O'doherty, H. Kaube, R. J. Dolan, and C. D. Frith (2004b): "Empathy for pain involves the affective but not sensory components of pain," *Science*, 303, 1157-1162.
- Singer, T., B. Seymour, J. P. O'doherty, K. E. Stephan, R. J. Dolan, and C. D. Frith (2006): "Empathic neural responses are modulated by the perceived fairness of others," *Nature*, 439, 466-469.
- Sobel, J. (2005): "Interdependent Preferences and Reciprocity," *Journal of Economic Literature*, 43(2), 392-436.
- Tangney, J.P., and R. Dearing (2002): *Shame and guilt*, New York: Guilford.
- Tvesky, A., and D. Kahneman (1992): "Advances in Prospect Theory: Cumulative Representaion of Uncertainty," *Journal of Risk and Uncertainty*, 5(4), 297-323.
- Xiao, H., and D. Houser (2005): "Emotion expression in human punishment behavior," *Proceedings of the National Academy of Sciences of the United States of America*, 102(20), 7398-7401.
- Xiao, H., and D. Houser (2009) "Avoiding the sharp tongue: Anticipated written messages promote fair economic exchange," *Journal of Economic Psychology*, 30(3), 393-404.