

La rationalité égoïste des préférences *altruistes*

Jean-François Grégoire, *Université Laval*

Le titre de ma conférence est «la rationalité égoïste des préférences altruistes», mais il aurait pu aussi bien être «la rationalité égoïste des préférences sociales» puisque je vais aborder les deux types d'explications des comportements pro-sociaux : 1-celle de la rationalité égoïste virtuelle de Philip Pettit et 2-celle des préférences sociales de Ernst Fehr et ses nombreux collaborateurs – particulièrement Fehr et Schmidt (2000). Le sujet de mon exposé porte sur l'explication des comportements pro-sociaux (comme, par exemple, la punition désintéressé par un tiers) dans les sciences sociales s'intéressant à la rationalité – et à la raison pratique – et aux déterminants des choix et des actions des agents. Je m'intéresserai plus particulièrement à l'enjeu concernant l'explication de la robustesse des normes sociales. Je présenterai la théorie de Pettit qui explique comment les normes sociales survivent dans une variété de contextes et ce, malgré la présence de comportements alternatifs disponibles aux agents.

Des chercheurs de différents horizons ont intérêt à bien cerner et comprendre les motivations et les déterminants de l'action pour produire ensuite un discours normatif, que ce soit en éthique, en théorie politique, en économie ou autre. La version de la rationalité (pratique) adoptée n'est pas sans importance dans un discours normatif. Elle a d'ailleurs souvent influencé le choix de principes de justice en théorie politique – pensons aux différents entre Harsanyi et Rawls par exemple. Or, avec l'utilisation de la méthode expérimentale en sciences sociales (incluant la philosophie), de nouvelles questions sont posées et de nouvelles méthodes voient le jour, mais l'enjeu philosophique de fond demeure le même : offrir des réponses aux questions normatives traditionnelles concernant les institutions et le «devoir-être».

Le but de ma communication est d'apporter une contribution philosophique aux débats interdisciplinaires fondamentaux qui ont actuellement cours dans les sciences sociales. La problématique que je soulève ici est celle présentée par Francesco Guala¹, à savoir la confusion générée par les interprétations de la robustesse de certains de modèles empiriques – en économie – qui sont utilisés comme arguments pour soutenir la validité externe des résultats expérimentaux dans l'explication des comportements pro-sociaux. «Indeed, robustness testing should be kept well separated from external validity... When this difference is not adequately appreciated, experimental debates tend to generate some confusion².» En d'autres mots, plusieurs expérimentateurs confondent la robustesse de leurs modèles en environnement contrôlé et la robustesse de leur théorie explicative des décisions et des actions dans une grande variété de contexte. La première explication n'est qu'une description d'un phénomène observé dans un contexte précis, alors que la seconde est la théorie générale que l'on désire posséder. L'enjeu est donc de dissiper la confusion entre les différents niveaux explicatifs en distinguant la robustesse des comportements en environnements contrôlés de la robustesse des comportements pro-sociaux dans une grande variété de contextes. On cherche à expliquer comment les comportements pro-sociaux (ou les normes sociales) se maintiennent dans le temps et résistent à la concurrence de comportements alternatifs, comme les comportements égoïstes (*free-riding*).

Je soutiens que la théorie du contrôle virtuel de la rationalité-égoïste de Philip Pettit³ nous permet d'éviter les confusions générées par les explications des comportements pro-sociaux formulées en termes de «préférences sociales» et de pouvoir resituer le rôle de la rationalité égoïste dans le raisonnement pratique. C'est ainsi que la théorie de Pettit nous permet d'accéder à un niveau de généralité assez élevé pour expliquer une multitude de phénomènes. C'est ce niveau de généralité qui nous permet de donner aux interprétations des données et des phénomènes une validité externe porteuse pour construire ensuite un discours normatif. L'objectif de cette démarche est de pouvoir expliquer la robustesse des normes dans une grande variété de contextes.

Pour présenter mon argument, je débiterai par présenter les théories des préférences sociales. Je soulignerai au passage deux failles dans ce type d'explication et m'intéresserai davantage à la seconde. Ensuite, je présenterai la théorie de la théorie du choix rationnel (TCR) de Pettit afin de montrer en quoi elle nous permet 1-de restituer la force explicative de la TCR et 2-de dépasser les explications en termes de « préférences sociales » pour fournir l'explication manquante nécessaire au passage à la validité externe. Se concentrer sur l'enjeu de l'explication de la robustesse des normes sociales nous permet de limiter la portée normative des modèles empiriques et d'adopter une perspective qui soit également normative, dans la mesure où il devient clair que les agents acceptent certaines contraintes – morales – sur leur comportement et agissent conformément à celles-ci dans une large variété de contextes d'interaction.

1. Les modèles des préférences sociales

Une des explications ayant connu le plus de succès en sciences sociales pour rendre compte des comportements pro-sociaux est celle des *préférences sociales*⁴. Les modèles qui soutiennent cette explication ont été notamment développés en économie expérimentale et cherchent à expliquer certaines observations empiriques contraires à la rationalité économique standard de l'égoïsme rationnel (*Homo Economicus*). Ce type de rationalité ressemble à celui du point de vue personnel chez Thomas Nagel en ce que tous deux les présentent comme étant en tension avec les préférences sociales ou le point de vue impersonnel⁵. À la différence de Nagel, les économistes expérimentaux ne cherchent pas à mettre en tension deux types de préférences, mais plutôt à fondre différents types de préférences dans une même fonction d'utilité, capable d'expliquer les comportements « altruistes ».

La méthode expérimentale peut nous informer sur les processus cognitifs des agents réels lorsqu'ils prennent des décisions, notamment lorsqu'ils se conforment à des normes sociales⁶. Le point crucial est de savoir si les théories des préférences sociales offertes par les économistes parviennent vraiment à nous faire avancer dans

une enquête sur la rationalité et les déterminants de l'action et, surtout, à expliquer la robustesse des comportements pro-sociaux. Avant d'introduire les deux failles dans les explications en termes de préférences sociales que je souhaite pointer ici, je crois qu'il est important de mentionner que les théories des préférences sociales arrivent à une conclusion somme toute assez triviale, à savoir que les individus ne sont pas uniquement préoccupés par la manière dont ils sont traités personnellement, mais aussi par le traitement réservé à autrui. Le point crucial est de savoir comment il est possible de réconcilier cette intuition avec notre autre intuition à propos du fait que les agents agissent parfois en vertu de considérations instrumentales et égoïstes.

1.2. Deux failles des interprétations en termes de préférences sociales

Je vais maintenant pointer les deux failles les plus saillantes des explications en termes de préférences sociales et montrer en quoi elles posent problème. Chacune d'elles mériteraient un traitement détaillé, mais je vais me concentrer ici sur la seconde.

La première faille importante des explications des comportements pro-sociaux en termes de préférences sociales réside dans les nombreuses confusions conceptuelles qui truffent ces théories⁷. Le problème trouve sa source dans l'arrière-plan flou de ces théories et dans leurs présupposés évolutionnistes. Comme le remarque justement Guala :

Economists rely on various informal accounts (typically, evolutionary or arbitrage stories)... to define these stories as «theoretical arguments» is somehow exaggerated, because they seldom take the form of rigorous theories... Experimenters are not looking for a theoretical explanation of the mechanism, which is therefore confined in the «background» of the theory⁸.

Deux éléments importants sont contenus dans cette citation. D'abord, les récits arbitraires et évolutionnistes qui sous-tendent les théories des économistes expérimentaux. Ensuite, le fait que les

expérimentateurs ne fournissent pas d'explication théorique et que ce niveau d'explication demeure implicite ou flou. Donc, la première faille tient dans le fait que les théories des préférences sociales comportent des confusions conceptuelles. Cela disqualifie d'emblée les explications des comportements pro-sociaux qui présentent cette faille. Une première contribution que les philosophes peuvent apporter est de faire un travail de clarification conceptuelle afin de rendre les théories plus rigoureuses.

La deuxième faille que je souhaite pointer ici, en soulignant sa grande importance pour mon propos, est celle que je nomme le « déficit de généralité » des théories des préférences sociales. Cela recoupe le second point que je viens de mentionner dans la citation de Guala quant au niveau d'explication désiré, à savoir une explication théorique – et robuste – des phénomènes observés empiriquement. D'abord, il faut distinguer trois types d'explications : 1-émergence des comportements et des normes, 2-leur maintien dans un contexte donné et 3-la robustesse des normes en général, c'est-à-dire leur robustesse dans une grande variété de contexte. Or, comme le démontre les développements en économie et en philosophie expérimentales, le déficit de généralité des théories des préférences sociales s'explique par le fait que ces théories ne prennent pas assez sérieusement en compte l'élément « contexte »⁹.

La structure d'interaction modifie considérablement la manière dont les agents prennent des décisions – et donc leur comportement. Or, les théories des préférences sociales ont été développées suite à la répétition d'expériences dans lesquelles les contextes d'interaction étaient très semblables – il s'agit en fait du jeu de l'Ultimatum¹⁰. D'autres chercheurs ont par la suite cherché à tester la robustesse de l'interprétation des modèles empiriques formulées en termes de préférences sociales, c'est le cas notamment de la philosophe Cristina Bicchieri. Cette dernière a démontré de manière convaincante que les théories des préférences sociales sont incapables d'expliquer des résultats obtenus suite à une modification – parfois mineure – de la structure d'interaction en contexte expérimental. Les conclusions de Bicchieri sont claires : les modèles de préférences sociales existant ne sont pas satisfaisants en raison de leur incapacité à expliquer

les comportements dans une large variété de contextes. C'est ce que je nomme la faille du «déficit de généralité» de ces théories des préférences sociales. Ce déficit s'explique par le fait que les économistes confondent l'explication empirique (modélisation) des comportements pro-sociaux avec l'explication théorique. Or, une théorie cherchant à expliquer la robustesse des comportements et des normes dans une grande variété de contextes doit fournir l'explication manquante, l'explication théorique.

Les deux failles que je viens de présenter peuvent être évitées par la théorie de Pettit qui permet d'expliquer les résultats empiriques obtenues dans différentes expériences, différents contextes et différentes cultures¹¹. C'est le niveau suffisant de généralité de la théorie qui permet d'expliquer ces variations et qui confère à la théorie de Pettit la validité externe dont la carence caractérise les théories des préférences sociales – qui ne sont finalement que des modèles. Je vais maintenant présenter la théorie de la réalité virtuelle de l'*Homo Economicus* de Pettit pour montrer en quoi elle nous permet – encore aujourd'hui – de dépasser le cadre dominant en économie pour penser la rationalité et les déterminants de l'action des agents.

2. La théorie de la réalité virtuelle de l'Homo Economicus chez Pettit

Je vais maintenant introduire la théorie de la réalité virtuelle d'*Homo Economicus* de Philip Pettit en présentant les grandes lignes de celle-ci. J'insisterai sur les points forts de la théorie qui nous permettent de pouvoir surmonter les failles qui caractérisent les théories des préférences sociales – malgré les avancements réalisés par les économistes, notamment grâce à l'approche expérimentale.

L'objectif de Pettit lorsqu'il aborde la question de la TCR est d'en préciser le rôle explicatif en tant que déterminant des décisions et actions des agents. Il s'agit en fait de théoriser rigoureusement l'intuition dominante en économie selon laquelle les agents sont généralement préoccupés par leurs intérêts égoïstes. Or, une fois cette intuition admise, il reste à savoir de quelle manière précisément l'égoïsme rationnel est opérant dans la prise de décisions des

individus et comment il détermine leurs actions. La thèse de Pettit est que la rationalité d'*Homo Economicus* opère de manière virtuelle sur le «pilote automatique culturel», c'est-à-dire sur les habitudes de socialisation et les préférences normatives individuelles. Il est important de souligner qu'en marge de l'objectif explicatif de Pettit, la théorie revêt un aspect normatif en ce qu'elle suggère que les agents agissent la grande majorité du temps dans le but de se conformer aux normes sociales. C'est uniquement lorsque la structure du contexte d'interaction ne satisfait pas les contraintes que l'agent admet habituellement sur son action que ce dernier sera tenté d'adopter une perspective conséquentialiste. La rationalité égoïste agit comme un système de contrôle virtuel sur nos habitudes de socialisation – un peu comme une automobile qui vous avertirait lorsque vous bifurqueriez de votre voie sans avoir clignoté; vous seriez causalement responsable de la conduite du véhicule, mais le mécanisme d'avertissement agirait comme une cause indirecte (*standby cause*). Je vais maintenant présenter une citation de Pettit pour rendre les choses plus limpides :

People may proceed under more or less automatic, cultural pilot in most cases but, at any point where a decision is liable to cost them dearly in self-regarding terms, the alarm bells ring and prompt them to consider personal advantage... they are disposed to do the *relatively* more self-regarding thing¹².

Quelques points importants émergent de cette citation. D'abord, le «pilote automatique culturel» – ou les habitudes de socialisation – explique la majorité des comportements des agents et ce, sans que la rationalité d'*Homo Economicus* intervienne. Or, nous pouvons considérer, comme le fait Pettit, que l'intérêt des agents peut être social – et non seulement économique – particulièrement lorsqu'ils suivent en contexte où des points de références, ou des normes, sont saillants. Je reviendrai plus en détail sur cette affirmation dans quelques instants. Pour le moment, le point que je cherche à souligner est simplement que la manière dont Pettit caractérise l'intérêt des individus est moins problématique que les histoires arbitraires ou évolutionnistes flous qui sous-tendent les théories des

préférences sociales. Le fait d'introduire des considérations pour les normes sociales dans l'analyse de la TCR et dans l'explication des comportements pro-sociaux nous permet d'éviter les confusions conceptuelles qui constituent la première faille des explications en termes de préférences sociales. La méthodologie et le langage de Pettit nous permettent d'éviter cette première faille dans notre explication des comportements pro-sociaux.

Ensuite, j'ai souligné dans la citation le fait que les individus dont les alarmes de la rationalité égoïste sont activées sont disposés à faire ce qui est «relativement» (*relatively*) dans leur intérêt. C'est là un point qui peut sembler flou ou encore trivial, mais qui n'est pas sans importance. La théorie de Pettit vise à restituer rigoureusement le rôle de la TCR dans l'explication des comportements pro-sociaux, afin de pouvoir fournir l'explication manquante, celle de la robustesse des normes. L'enjeu est de fournir une théorie qui soit capable d'expliquer pourquoi les agents suivent de manière robuste les normes sociales malgré qu'ils soient parfois tentés de satisfaire leurs intérêts égoïstes en adoptant une perspective instrumentale – et malgré qu'ils soient conscients de cette possibilité et du fait que d'autres l'actualisent. Malgré que les agents dévient parfois de leurs habitudes comportementales, la robustesse des comportements pro-sociaux s'explique par le fait que dans la majorité des cas le contexte d'interaction génère des attentes et des contraintes que l'agent est prêt à accepter et qui détermineront donc – en partie – son action.

Puis, je crois que le pilote automatique culturel franchit les portes du laboratoire lorsque les économistes réalisent leurs expériences. Malgré que la manière dont les interactions sont structurées ait un impact sur les décisions des agents (effet de «cadrage»), les points de références en matière d'équité – ou autre – découlant des habitudes de socialisation entrent dans le laboratoire. Je soutiens que si l'on prend en compte l'effet de cadrage, nous pouvons interpréter les résultats expérimentaux et les phénomènes qui en émergent à l'aide de la théorie de Pettit. Nous devons être sensibles aux contextes d'interactions et voir lesquels sont susceptibles d'activer les alarmes de la TCR¹³. L'anonymat en laboratoire n'est pas suffisant pour dissocier les points de références en matière d'acceptabilité sociale

du désir rationnel et égoïste d'acceptabilité sociale des agents. Ce que Pettit désigne par l'intérêt social des individus est en fait ce désir rationnel et égoïste d'acceptabilité social. Or, si ce désir franchit les portes du laboratoire conjointement avec les points de références culturellement (pré-)déterminés, nous pouvons voir en quoi la théorie de Pettit nous permet d'avoir une théorie capable d'expliquer les comportements des agents dans une grande variété de contextes. Les comportements pro-sociaux, comme la punition par les tiers, sont souvent motivés par une émotion liée à l'insatisfaction des attentes des agents. Dans la majorité des cas, les agents acceptent de répondre aux attentes raisonnables d'autrui. Dans les cas limites, le contexte favorise le déclenchement des alarmes de l'égoïsme.

Il s'agit d'expliquer la robustesse des comportements et des normes au niveau théorique, et non uniquement leur émergence ou leur robustesse dans des environnements contrôlés. Les différents types d'explication ne sont pas en concurrence, ils se complètent. Seulement, l'explication de la robustesse des normes de Pettit nous permet de passer de la modélisation à la théorisation en déployant une théorie qui demeure habituellement dans l'arrière-plan flou des modèles empiriques.

2.1. La virtualité de la TCR chez Pettit et le passage à la validité externe

Les déterminants culturels façonnent le comportement et peuvent expliquer l'émergence et/ou le maintien dans le temps de certains comportements altruistes (comme la punition par les tiers). Mais l'explication de la robustesse des comportements pro-sociaux hors du laboratoire – dans la vie de tous les jours – est différente. Le point crucial est d'expliquer pourquoi les agents adoptent des comportements pro-sociaux dans une grande variété de contextes, malgré la présence de comportements alternatifs, notamment les comportements explicitement égoïstes – pour lesquels la rationalité égoïste agit comme une cause directe, i.e. ce que l'on nomme familièrement le comportement du resquilleur ou du passager clandestin (*free-rider*). C'est uniquement en déployant ce genre d'explication que nous pouvons arriver à une théorie assez générale

pour expliquer les comportements des agents dans un grande variété de contextes et, surtout, obtenir une validité externe à partir de ces explications, c'est-à-dire pouvoir passer de la modélisation à la théorisation.

L'explication que fournit Pettit nous permet d'expliquer quelque chose *de plus* que les interprétations courantes des expérimentalistes. Le fait que la théorie de Pettit soit sensible aux contextes d'interactions et à la socialisation des individus lui donne le niveau de généralité qui manque aux modèles empiriques des économistes. La théorie de Pettit peut rendre compte des variations interculturelles et de la robustesse intra-culturelle des résultats expérimentaux. Le problème que j'ai présenté en introduction en m'appuyant sur la distinction de Guala entre la robustesse des modèles ou des explications empiriques et la validité externe des explications qui leur donne le niveau théorique recherché peut être résolu à l'aide la théorie de Pettit. La TCR joue un rôle virtuel au niveau social en plus de jouer un rôle de premier plan au niveau économique – là où les contextes favorisent le déclenchement des alarmes de la TCR. Si l'on ne comprend pas l'intérêt égoïste comme étant limité à la maximisation économique ou monétaire, mais que l'on introduit la dimension sociale qui caractérise la normativité en général ainsi que l'importance du contexte (révélée par les avancements récents dans les domaines expérimentaux des sciences sociales), nous pouvons dépasser le cadre des théories des préférences sociales et développer un cadre plus général et plus performant. L'aspect normatif de la théorie de Pettit nous permet d'expliquer la robustesse des comportements par le désir des agents de se conformer aux normes sociales. La théorie de Pettit nous permet également d'éviter le galvaudage dans le passage de la modélisation à la théorisation. L'argument général de ma communication peut donc se résumer ainsi : la philosophie peut grandement contribuer aux débats interdisciplinaires en sciences sociales en apportant (une précision conceptuelle et) une théorie de la raison pratique qui puisse nous permettre de bien utiliser les modèles empiriques – très utiles pour tester certaines hypothèses et l'influence de certaines variables bien précises. L'argument plus spécifique est d'ajouter un troisième *explanandum* aux explications des comportements pro-sociaux : non

seulement l'émergence et la « répétition » de ceux-ci dans certains contextes, mais la robustesse des actions et des points de références culturellement (pré-)fixés – i.e. des normes, car ce sont elles qui ont un véritable intérêt du point de vue normatif.

Selon l'argument de Pettit que je présente ici, la robustesse des comportements pro-sociaux (comme la punition par les tiers) s'explique – en grande partie – par le désir égoïste des agents à satisfaire leur désir virtuel d'acceptabilité sociale et par leur désir quasi « automatique » de respecter les points de références culturellement fixés. Ce désir virtuel agit comme une cause indirecte dans les comportements pro-sociaux, mais doit tout de même être pris en compte afin de ne pas être limité au niveau de la modélisation et de la description des phénomènes. Cet aspect normatif de la théorie est important. On cherche à pouvoir expliquer les phénomènes et à fournir l'explication théorique qui demeure généralement implicite ou laissée pour compte par les théories économiques de la rationalité. La théorie de Pettit nous permet d'y arriver et de préciser rigoureusement comment opère la rationalité égoïste des agents – plutôt que de limiter celle-ci à une intuition directrice. Concrètement, l'aspect normatif de la théorie nous permet de surmonter le déficit de généralité des théories des préférences sociales en affirmant que les points de référence en matière d'acceptabilité sociale guidant le pilote automatique culturel sont ce vers quoi la TCR oriente virtuellement les agents de manière robuste; et ce, à l'intérieur ainsi qu'à l'extérieur du laboratoire.

Conclusion

La réflexion philosophique sur le raisonnement pratique et les déterminants des décisions et actions des agents peut apporter beaucoup aux sciences sociales qui cherchent à préciser leurs théories de la décision dans le but d'orienter conséquemment leurs théories normatives. Les théories des préférences sociales avancées par certains expérimentateurs ont permis de mettre en lumière certains éléments importants. Toutefois, elles n'offrent pas un cadre adéquat pour les chercheurs visant à fournir des réponses aux questions normatives traditionnelles concernant les institutions et le « devoir-être » de

manière général. Dans la présente communication, j'ai cherché à démontrer la possibilité et la nécessité d'apporter une contribution philosophique à ces enjeux. D'abord, les philosophes doivent poursuivre leur tâche par excellence et clarifier certains concepts fondamentaux qui sont utilisés dans différents débats dont les enjeux débordent les frontières de la discipline. Ensuite, les philosophes doivent apporter une contribution aux débats interdisciplinaires en informant les autres disciplines à propos des réflexions portant sur le raisonnement pratique et les déterminants de l'action. J'ai présenté ici la théorie de la réalité virtuelle d'*Homo Economicus* de Pettit comme un exemple percutant de ce type de contribution aux explications des comportements pro-sociaux. J'espère vous avoir convaincu non pas de la théorie de Pettit qui mériterait d'être discutée plus en détail, mais de l'importance de distinguer les différents niveaux d'explication des comportements pro-sociaux afin de voir où se situe l'enjeu philosophique et explicatif. Cette introduction d'un aspect normatif dans l'explication de la robustesse des comportements pro-sociaux est utile pour créer des ponts entre les différentes disciplines s'intéressant au raisonnement pratique et aux déterminants de l'action dans le but de penser les institutions.

-
1. Francesco Guala. *The methodology of experimental economics*, Harvard University Press, Boston MA, 2005.
 2. *Ibid*, p. 225.
 3. Philip Pettit. «The virtual reality of *Homo Economicus*», *Rules, Reason and Norms*, New York, Oxford University Press, 2002.
 4. Pour une revue récente des différents modèles des préférences sociales disponibles sur le marché, voir Dinky Daruvala, «Would the right social preference model please stand up!» dans *Journal of Economic Behavior and Organisation*, no. 73, 2010, p. 199-208.
 5. Thomas Nagel, *Égalité et partialité*, Paris, Presses Universitaires de France, 1994.
 6. Pour un exemple éclairant de ce genre de travail, voir Cristina Bicchieri, *The Grammar of Society*, New York, Oxford University Press, 2006.
 7. Pour un exemple percutant de ce genre de critique, voir Mark S. Peacock, «The conceptual construction of altruism: Ernst Fehr's experimental

- approach to human conduct» dans *Philosophy of the Social Sciences*, vol.37, no. 1, 2007, p. 3-23.
8. Francesco Guala, *op. cit.*, p. 226-227.
 9. Il existe une vaste littérature sur ce qui a été nommé «l'effet de cadrage» (*framing effect*). Pour des exemples majeurs, voir le papier phare de Amos Tversky et Daniel Kahneman, «The framing of decisions and the psychology of choice», *Science*, vol. 211, 4481, 1981, p. 453-458; également, l'ouvrage de la philosophe Cristina Bicchieri, *The Grammar of Society*, Oxford University Press, New York, 2006.
 10. Guala fait la distinction entre «repetition» et «replication» pour questionner la validité externe des modèles empiriques. *Op. cit.*, p 14.
 11. Pour les variations interculturelles, je me réfère ici aux travaux du méga projet interdisciplinaire lancé par l'anthropologue Joseph Henrich. Voir notamment *Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scales societies*, J. Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr (eds), Oxford University Press, Oxford, 2004.
 12. Philip Pettit. «The Virtual Reality of *homo economicus*», *Rules, Reason and Norms*, New York, Oxford University Press, 2002 (je souligne).
 13. Pour un exemple intéressant de contexte où ces alarmes sont activées, voir Dana, J., Cain, D.M., and Dawes, R. «What you don't know won't hurt me: Costly (but quiet) exit in a dictator game », dans *Organizational Behavior and Human Decision Processes*, vol. 100, no. 2, 2006, p. 193-201.