

ÉMOTIONS MORALES ET COMPORTEMENT PROSOCIAL : UNE REVUE DE LA LITTÉRATURE

Romain Jourdheuil, Emmanuel Petit

Dalloz | « Revue d'économie politique »

2015/4 Vol. 125 | pages 499 à 525

ISSN 0373-2630

Article disponible en ligne à l'adresse :

<http://www.cairn.info/revue-d-economie-politique-2015-4-page-499.htm>

!Pour citer cet article :

Romain Jourdheuil, Emmanuel Petit, « Émotions morales et comportement prosocial : Une revue de la littérature », *Revue d'économie politique* 2015/4 (Vol. 125), p. 499-525.

DOI 10.3917/redp.254.0499

Distribution électronique Cairn.info pour Dalloz.

© Dalloz. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

Émotions morales et comportement prosocial : Une revue de la littérature

Romain Jourdheuil*

Emmanuel Petit*

• BILAN / ESSAI

Dans cet article, nous proposons une revue de la littérature consacrée au rôle spécifique des émotions morales en économie. Nous nous appuyons sur les résultats empiriques issus des travaux en économie expérimentale et en psychologie pour analyser l'impact des émotions morales sur l'émergence des comportements prosociaux comme la coopération, l'altruisme, la confiance et la réciprocité. La sanction affective, liée à l'influence des émotions morales négatives (culpabilité, honte, colère...), conduit à la régulation des comportements opportunistes tandis que la récompense affective, issue de l'émergence des émotions morales positives (fierté, gratitude...), favorise l'adoption de comportements socialement ou moralement valorisés. Nous analysons ensuite la façon dont s'effectue la prise en compte formelle du rôle des émotions morales en économie comportementale. Au sein des modèles de « préférences sociales », les émotions sont intégrées, d'une façon minimaliste, sous la forme de préférences exogènes qui sont activées parce que préexistent implicitement des normes sociales et morales contraignant les choix des individus. En tenant compte d'une conception élargie et plus réaliste des affects, l'impact des émotions sur les comportements dépend des intentions des agents, de leurs attentes, de leur personnalité ainsi que du contexte dans lequel s'effectue le processus de décision, ce qui conduit à considérer les émotions morales comme des préférences endogènes.

émotion morale – préférences sociales – économie expérimentale – économie comportementale

Moral emotions and prosocial behaviors: A review

This article provides a literature review on the role of moral emotions in economics. Drawing on empirical results obtained in experimental economics and psychology, we first discuss the impact of moral emotions on the emergence of other-regarding behaviors such as cooperation, altruism, trust and reciprocity. We especially highlight how the affective sanction associated with negative moral emotions (guilt, shame, righteous

* Université de Bordeaux, GREThA, UMR CNRS 5113, Avenue Léon Duguit, 33608 Pessac.
Courriel : romain.jourdheuil@u-bordeaux.fr

* Université de Bordeaux, GREThA, UMR CNRS 5113, Avenue Léon Duguit, 33608 Pessac.
Courriel : emmanuel.petit@u-bordeaux.fr ; Site Internet : <http://ecopsycho.gretha.cnrs.fr/>

anger...) participates in the regulation of opportunistic behaviors, whereas the affective reward derived from positive moral emotions (pride, gratitude...) promotes socially or morally valued behaviors. Second, we focus on how economists have responded to the challenge of integrating a role for moral emotions into economic theory. We particularly point out that, in a minimalist conception, emotions can be formalized as exogenous preferences activated because of the behavioral constraints embodied in pre-existing implicit social or moral norms. Nevertheless, a broader and (we argue) more realistic conception of emotions would allow their influence to depend on agents' intentions, expectations, and personalities as well as on the specific context of a decision, implying that moral emotions may be better understood as a form of endogenous preferences.

moral emotions – other-regarding preferences – experimental economics – behavioral economics

Classification JEL: C70, C90, D03, A12

1. Introduction

Depuis le travail séminal de Franck [1988] et la revue de la littérature proposée par Elster [1998] dix ans plus tard, les travaux portant sur le rôle des affects en économie ont connu une croissance rapide. Sous de multiples formes – émotions, sentiments, tempéraments, humeurs – les affects ont été plus largement intégrés aussi bien en économie expérimentale que comportementale dans les différents champs de l'analyse (Rick et Loewenstein [2008], Petit [2009]). Il existe par exemple une littérature innovante sur l'effet inconscient des humeurs ou des tempéraments sur les comportements individuels (Kirchsteiger, Rigotti et Rustichini [2006]). De même, les émotions contrefactuelles, comme le regret ou la déception, ont été abondamment mobilisées dans l'étude des comportements risqués (Giorgetta *et al.* [2013]).

L'un des intérêts de cette littérature récente est de prendre en compte la diversité et la spécificité du domaine affectif. En ce sens, un travail important de différenciation des affects a été effectué dans les études menées par les économistes : au lieu de considérer l'affect comme un concept englobant et général, les analyses se focalisent davantage sur l'impact d'une émotion spécifique (le regret, la colère, l'envie, etc.), sur l'effet d'une humeur positive ou négative ou encore sur les conséquences comportementales associées à une classe d'émotions particulière.

Au sein de cette vaste littérature, l'étude du rôle des « émotions morales » en économie revêt une dimension particulière. En se focalisant en effet sur ce type d'émotions, les économistes renouent avec la tradition des Lumières

écossaises (David Hume [1739], Adam Smith [1759]), interrogeant à nouveau le rôle des « sentiments moraux » dans l'émergence des comportements prosociaux. Les émotions morales – comme la culpabilité, la honte, l'indignation ou la gratitude – ont ceci de particulier qu'elles sont en lien avec la façon dont sont perçues les normes sociales ou morales au sein de la société. En règle générale, une émotion morale peut survenir lorsqu'une norme est respectée ou, au contraire, violée. L'émotion morale peut donc dans ce cas être à l'origine des comportements individuels de coopération, d'altruisme, de confiance, de réciprocité ou d'équité, que de nombreux travaux en économie comportementale ont mis en évidence (Camerer [2003]). Par exemple, dans un dilemme social, comme celui du célèbre dilemme du prisonnier, il est possible d'instaurer la coopération entre les acteurs si l'on suppose une capacité empathique des individus ou encore une propension à la culpabilité.

Dans cet article, nous proposons une revue de la littérature non exhaustive permettant de cerner le rôle spécifique des émotions morales en économie. Nous étudions, dans un premier temps, les preuves expérimentales mettant en évidence l'impact des émotions morales positives et négatives sur le comportement prosocial individuel (2). Nous analysons ensuite comment les émotions morales peuvent être introduites dans la modélisation économique (3). Nous concluons enfin brièvement (4).

2. L'impact des émotions morales sur les comportements prosociaux

Pour analyser l'impact des émotions morales sur les comportements, nous étudions tout d'abord la spécificité des émotions morales (2.1). Nous distinguons ensuite les effets des émotions négatives, via la sanction (2.2), de ceux induits par les émotions positives, via la récompense (2.3).

2.1. Qu'est-ce qu'une émotion morale ?

Dans un article séminal, Elster [1998] propose une typologie des différentes émotions au sein de laquelle on trouve le groupe des émotions dites « sociales ». Les émotions sociales comprennent la colère, la haine, la fierté, l'orgueil, la culpabilité, la honte, l'admiration ou l'affection. Ces émotions peuvent être tournées vers soi ou dirigées vers autrui (son comportement ou sa personnalité) ; leur valence peut être positive ou négative. Les émotions morales forment une sous-catégorie au sein du groupe des émotions sociales et ont par conséquent pour caractéristique centrale le lien avec autrui et une dimension sociale.

Pour obtenir une définition plus précise de cette famille d'émotions, il faut se tourner cependant vers la psychologie et en particulier en direction des travaux du psychologue social Jonathan Haidt. D'après Haidt [2003, p. 276, notre traduction], les émotions morales sont celles qui sont « *intrinsèquement liées aux intérêts et au bien-être de la société dans son ensemble, ou au moins au bien-être d'un autre individu que l'agent qui les ressent* ». Cette définition est, à première vue, assez générale. Haidt [2003] la complète en précisant qu'une émotion morale comporte deux dimensions spécifiques. La première dimension évalue les conditions d'émergence de l'émotion et, en particulier, dans quelle mesure l'évènement déclencheur de l'émotion est en lien avec l'intérêt individuel du sujet qui la ressent plutôt qu'avec celui d'un ou plusieurs autres agents. La deuxième dimension concerne l'action qui est motivée par l'émotion morale : dans quelle mesure la « tendance à l'action » pousse l'individu qui la ressent à agir dans son propre intérêt, plutôt que dans l'intérêt d'un ou plusieurs autres agents. Ainsi, plus une émotion peut être qualifiée de « désintéressée », par rapport à ces deux dimensions, plus celle-ci peut être considérée comme morale. Les émotions morales ont également comme caractéristiques d'apparaître lorsqu'une violation morale est perçue par les individus. Ainsi, comme l'atteste de nombreux travaux en psychologie expérimentale, dans de nombreuses situations d'interaction sociale, plus que notre capacité de raisonnement moral, ce seraient nos intuitions morales, en partie à travers le ressenti d'émotions morales, qui conduiraient à l'élaboration de notre jugement moral et qui, en conséquence, influeraient sur notre conduite (Haidt [2001], Vélez Garcia et Ostrosky-Solis [2006]).

Haidt [2003] propose une typologie des émotions morales qui comprend les émotions auto conscientes, les émotions d'évaluation d'autrui et l'empathie (voir le Tableau 1 ci-dessous). Pour chaque famille, il est ainsi possible de distinguer la valence de l'émotion (le plaisir ou le déplaisir qui y est associé), le déclencheur (le type de cognition ou de croyance à l'origine de l'affect) ainsi que la « tendance à l'action » propre à l'émotion (Frijda [1986]).

Les émotions auto-conscientes sont celles de la honte, de l'embarras, de la culpabilité et de la fierté. Elles ont en commun d'être déclenchées par les événements interpersonnels négatifs ou positifs dont la cause est attribuée à soi (Tangney *et al.* [2007]). Il est, par exemple, possible de ressentir de la culpabilité ou de la honte, après s'être comporté de façon opportuniste en profitant du fruit d'un effort collectif sans avoir contribué à sa réalisation, ou au contraire de ressentir de la fierté à la suite d'un réel investissement dans la réalisation d'un projet collectif au sein de l'entreprise. Ces émotions constituent un indicateur précieux – un « baromètre moral » – qui nous informe sur l'acceptabilité sociale ou morale de nos conduites, en agissant comme des punitions ou des récompenses affectives. En particulier, lorsqu'un individu transgresse une norme morale, il peut ressentir de la culpabilité, de l'embarras ou de la honte, qui sont à l'origine d'un désagrément. *A contrario*, un comportement socialement ou moralement adapté suscite des sentiments positifs de fierté ou d'approbation de soi. D'après les psychologues, il peut exister des différences sensibles au sein de cette classe d'émotions : la honte, par exemple, induit un comportement de dissimulation alors que la tendance à l'action de la culpabilité est la réparation. Il semble également que le sujet honteux soit plus sensible au degré d'expo-

Tableau 1. Caractéristiques des principales émotions morales

Famille	Valence	Émotion	Déclencheur	Tendance à l'action
Émotions auto-conscientes	-	Honte et Embarras	Violation des normes sociales ou morales	Dissimulation, fuite
	-	Culpabilité		Réparation, action prosociale
	+	Fierté	Investissement	Renforcement
Émotions d'évaluation d'autrui	-	Colère, Indignation	Injustice	Vengeance, demande de réparation
	-	Dégoût	Perception d'individus se rabaissant	Évitement, condamnation
	-	Mépris	Sentiment de supériorité morale	Moquerie, irrespect
	+	Gratitude	Perception d'une action bénéfique pour soi, volontaire et intentionnelle	Réciprocité positive
	+	Respect et Admiration	Beauté morale, charité, gratuité	Aspiration à devenir un être moral
Émotions de souffrance d'autrui		Compassion et Empathie	Perception de la souffrance ou de la tristesse d'autrui	Aide, réconfort

sition (publique) de son acte. Comparativement, la culpabilité semble davantage réservée à la sphère privée et à la conscience du sujet.

Les émotions d'évaluation d'autrui se focalisent sur les comportements adoptés par les individus qui nous entourent. On distingue (i) les émotions à valence négative – comme la colère, l'indignation, le mépris et le dégoût – qui conduisent à la vengeance, à l'humiliation ou à l'ostracisme, et incitent à des mesures punitives – (ii) des émotions à valence positive – comme la gratitude ou l'admiration – qui sont notamment à l'origine d'une réciprocité positive récompensant l'adoption de comportements prosociaux.

Dans la typologie donnée par Haidt [2003], l'émotion d'empathie occupe une place particulière : il s'agit de la capacité à ressentir une ou des émotions, sur la base de la perception d'une ou de certaines émotions similaires ressenties par une autre personne. Davantage qu'une émotion purement discrète, l'empathie est un processus émotionnel à travers lequel les individus sont amenés à ressentir des émotions telles que la colère, la tristesse ou la joie en fonction des émotions qu'ils perçoivent chez une autre personne. Ce processus conduit généralement les personnes « empathiques » à aider, à reconforter et à tenter de diminuer la souffrance de l'autre. Il permet également de renforcer l'efficacité des différents mécanismes de punitions et de récompenses associés aux autres émotions morales.

2.2. La régulation des comportements : le rôle de la sanction affective

2.2.1. Les effets prosociaux de la colère et du dégoût

En décrivant un mécanisme de « punition altruiste », Fehr et Gächter [2000] ont montré, dans un jeu du bien public répété¹, que l'existence d'une possibilité de punition – coûteuse pour celui qui la met en œuvre – augmente significativement le niveau de la coopération au sein du groupe. La coopération est stimulée lorsque les contributeurs sanctionnent financièrement les non contributeurs sous l'effet d'une émotion négative, possiblement la colère ou l'indignation. Cet effet peut être anticipé par les joueurs opportunistes qui augmentent immédiatement leur niveau de contribution avant même la mise en place effective du mécanisme de sanction. L'affect négatif anticipé fonctionne ici comme une menace crédible (Frank [1988], Peeters et Vorsatz [2013]) et permet également d'éviter la propagation des comportements de passagers clandestins que l'on observe en l'absence de ce mécanisme lorsque le jeu est répété. Des mécanismes similaires ont été identifiés par Masclet *et al.* [2003] qui observent que la coopération peut être

1. Dans ce jeu, chacun des quatre joueurs reçoit une dotation initiale qu'il peut soit conserver (partiellement ou intégralement), soit allouer à une cagnotte commune (le bien public). Comme dans tous les dilemmes sociaux, la stratégie dominante de chaque joueur est de ne jamais contribuer alors qu'*a contrario* les paiements agrégés sont maximisés lorsque chaque joueur contribue l'intégralité de sa dotation.

soutenue par une simple désapprobation morale du comportement observé des autres joueurs. Carpenter *et al.* [2004] montrent aussi, qu'en milieu naturel, de nombreux participants sont prêts à payer pour exprimer leur désapprobation (au moyen d'une image représentant un visage mécontent) à ceux qui se montrent non coopératifs, sans nécessairement leur imposer une sanction matérielle.

Dans ces différents travaux, les émotions morales négatives permettent d'interpréter les réactions comportementales des sujets. Aucune mesure des émotions n'est cependant effectuée expérimentalement. Les travaux portant sur les jeux de négociation² ont permis d'affiner les liens entre le ressenti émotionnel des sujets et leur comportement observé. Bosman et van Winden [2002], Xiao et Houser [2005] et Ben-Shakhar *et al.* [2007] ont ainsi révélé le rôle dissuasif et efficace de la colère ou du mépris lorsque l'un des partenaires fait une proposition d'offre perçue comme injuste et intentionnelle. La colère conduit ici au rejet de l'offre et à la rupture de la négociation. Lorsque les sujets ont cependant l'opportunité d'exprimer leurs émotions de manière explicite (sur papier), les rejets sont moins fréquents, le message écrit s'interprétant comme une punition psychologique susceptible de faire naître de la honte ou de la culpabilité chez l'offreur. L'expression anticipée de la colère peut également avoir une influence directe sur l'offre du négociateur (Xiao et Houser [2009]). Lorsque, dans un jeu du dictateur, celui-ci s'attend à recevoir l'approbation ou la désapprobation (écrite) du destinataire de son offre, il a tendance à renoncer à capter l'intégralité du surplus. Moretti et di Pellegrino [2010] retrouvent ce type de mécanismes en s'intéressant à l'émotion de dégoût. Leur étude montre surtout la proximité entre la colère et le dégoût, deux émotions associées, entre autres, à l'activation d'une région cérébrale commune – l'insula antérieure (Sanfey *et al.* [2003]). Les auteurs supposent et confirment qu'une induction du dégoût provoque une tendance accentuée au rejet d'une offre profitable dans le jeu de l'ultimatum, la violation de la norme d'équité et le dégoût viscéral impliquant des mécanismes neuraux qui se recoupent.

Le respect des normes sociales peut également être obtenu via la participation d'un tiers – une personne qui ne subit pas le comportement immoral d'autrui mais qui en est le témoin. D'après Fehr et Fischbacher [2004], la plupart des normes morales ne seraient en fait pas préservées si seules les personnes directement concernées imposaient des sanctions, que l'on pense au vol, à la fraude, au mensonge ou à la malhonnêteté. Dans un jeu du bien public similaire à celui présenté par Fehr et Gächter [2000], Carpenter et Matthews [2012] observent en particulier que les membres d'un groupe sont susceptibles de punir à leurs frais des membres d'un autre groupe, dont ils observent le comportement mais avec lesquels ils ne jouent pas. Les niveaux de punitions constatés sont moins élevés que lorsque les individus sanctionnent les faibles contributeurs de leur propre groupe. En revanche, c'est bien l'introduction d'une possibilité de punition par une

2. Les jeux de négociation sont ceux de l'ultimatum ou du « pouvoir de prendre » (« *power-to-take game* »), dans lesquels l'un des joueurs fait une offre que le second peut accepter ou refuser, ou encore celui du « dictateur » dans lequel le receveur ne dispose plus de droit de veto (pour une revue de ces jeux, voir, respectivement, Petit [2010] et Petit et Rouillon [2010]).

tierce personne (non impliquée), associée à une punition plus directe des partenaires du groupe, qui permet d'obtenir une augmentation significative de la contribution des joueurs.

En déclenchant des représailles, la colère et le dégoût ont un impact direct sur le comportement de la victime ou du tiers. Ces émotions sont ici l'expression d'une réciprocité négative formalisée par Rabin [1993]. Ces émotions de condamnation d'autrui peuvent également déclencher d'autres processus émotionnels chez le transgresseur, qui l'inciteront à ne pas enfreindre la règle sociale en cours. C'est le cas notamment lorsque celui qui est le sujet moral d'une juste indignation, ressent de la culpabilité, du remords ou de la honte, qui l'obligent à reconsidérer son comportement initial. Le maintien de la norme provient ici à la fois d'une régulation externe (la victime ou la personne tiers) et d'une régulation interne ou intériorisée, via une émotion auto-consciente. L'étude de Reuben et van Winden [2010] confirme ce scénario dans un jeu du « pouvoir de prendre ». Les auteurs montrent en particulier que les offreurs qui sont l'objet d'une sanction à la suite d'une offre inéquitable initiale et qui en ressentent de la culpabilité, font ultérieurement une offre plus équitable. L'impact conjoint de la colère et de la culpabilité favorise donc l'émergence d'un comportement social altruiste.

2.2.2. Le rôle social des émotions anticipées : embarras, culpabilité et honte

D'après Frank [1988], les émotions morales auto-conscientes à valence négative – comme la culpabilité ou la honte – incitent les individus à choisir les comportements prosociaux qui possèdent un intérêt stratégique à long terme. Elles s'opposent à l'intérêt personnel, davantage orienté vers les récompenses immédiates (défection, trahison, procrastination), en rendant saillantes les conséquences émotionnelles négatives de comportements opportunistes. De tels agissements suscitent en effet des émotions déplaisantes (culpabilité, honte, embarras) qui peuvent être anticipées et rationnellement intégrées par l'individu dans sa fonction d'utilité de long terme. En régulant à long terme les comportements individuels, ces émotions ont ainsi un rôle crucial dans la construction et le maintien des relations économiques interpersonnelles.

Les expériences conduites par Brown [1970] illustrent bien ce mécanisme d'anticipation à l'œuvre en présence d'interaction sociale. Au cours de l'expérimentation, les participants ont le choix de plusieurs supports médias pour présenter la description d'une tâche de perception sensorielle qui peut être, selon la condition à laquelle ils appartiennent, embarrassante ou neutre. La rémunération perçue pour cette description dépend positivement du niveau d'exposition du média choisi. Classiquement, un individu rationnel doit se porter vers le support où le niveau d'exposition est le plus élevé, ce que ne confirment pas les résultats de l'étude. En particulier, les sujets de la condition « Embarras » sacrifient significativement plus d'argent que les individus de la condition « Contrôle », indiquant qu'ils sont davantage prêts à faire des sacrifices financiers pour ne pas ternir l'image de soi qu'ils donnent aux autres dans une situation publiquement embarrassante.

Les travaux plus récents de Ketelaar et Au [2003] ou de Reuben et van Winden [2010] valident empiriquement l'argument proposé initialement par Frank [1988]. Ketelaar et Au [2003] montrent, dans un dilemme du prisonnier répété, que l'induction du sentiment de culpabilité motive le comportement coopératif, notamment de ceux qui ont initialement transgressé la norme. La culpabilité a également un effet d'apprentissage visible dans le temps. Ainsi, dans un jeu de l'ultimatum, des participants ayant rapporté un sentiment de culpabilité après avoir effectué une première offre inéquitable, modifient leur stratégie une semaine plus tard en proposant une offre plus généreuse lors de la répétition du jeu. La culpabilité remplit donc un rôle d'information (consciente ou non) concernant le coût à long terme du non respect des normes sociales (comme l'équité ou la coopération). Ce signal de long terme est pris en compte par l'individu au moment de la prise de décision et peut conduire à l'influencer.

L'impact de la honte en tant que signal permettant de mobiliser à long terme des comportements prosociaux est moins évident. Il existe effectivement une abondante littérature expérimentale montrant que les individus ont tendance à adopter des comportements altruistes lorsqu'ils sont observés publiquement, ces comportements étant beaucoup plus limités dans la sphère privée (Petit et Rouillon [2010]). Ceci sous-tend que le comportement de don serait davantage issu d'un désir d'apparaître comme équitable plutôt que d'une véritable préférence pour l'équité. Certains travaux théoriques en ont déduit que le sentiment de honte, associé au désir de ne pas apparaître comme égoïste, pouvait motiver le comportement de don (Neilson [2009], Dessi et Zhao [2010]). Pour autant, sur le plan expérimental, les effets de la honte induite sur les comportements prosociaux sont mitigés. Hooge *et al.* [2007], utilisant un protocole similaire à celui de Ketelaar et Au [2003], trouvent ainsi un effet significatif de la culpabilité, mais aucun impact de la honte. La honte serait bien ainsi cette émotion « détestable » qui, fragilisant l'estime de soi de l'individu, l'incite à fuir la relation sociale plutôt qu'à la restaurer, contrairement à ce que la culpabilité est capable de faire. Le travail plus récent d'Hooge, Breugelmans et Zeelenberg [2008] remet en question cette vision univoque. De façon très intéressante, cette étude révèle que l'induction de la honte favorise effectivement la coopération mais uniquement dans certaines conditions. Pour avoir un impact, il faut, d'une part, que le comportement d'aide ou de coopération soit destiné à celui qui a été témoin de la situation honteuse. Dans ce cas, l'influence de la honte est dite endogène. Il faut, d'autre part, que la honte touche les individus dont l'orientation sociale est à l'origine individualiste et non prosociale. Comme dans l'étude de Ketelaar et Au [2003], ces résultats indiquent que l'influence des émotions s'effectue de manière différente selon la personnalité des individus, et en particulier en fonction de leur propension à ressentir de la culpabilité ou de leur degré d'orientation sociale.

2.2.3. Représailles, coût émotionnel et effets contreproductifs

Lorsqu'elles déclenchent des réactions de la part de leurs émetteurs, les émotions morales, comme la colère ou le dégoût, peuvent avoir en retour,

comme nous l'avons souligné, des effets bénéfiques pour la collectivité. De même, d'autres émotions morales, comme la honte ou la culpabilité, sont capables de limiter l'adoption de comportements déviants lorsque celles-ci sont anticipées par ceux qui s'apprêtent à violer la norme sociale ou morale.

Les émotions morales à valence négative ont néanmoins certaines limites qui ont également été mises à jour expérimentalement :

(i) Les émotions de colère, d'indignation, de mépris ou de dégoût peuvent tout d'abord se révéler contreproductives en suscitant à leur tour des émotions négatives chez l'individu qui a enfreint la norme collective. Ceci implique un risque de représailles, d'escalade du conflit et de blocage des négociations (Pietroni *et al.* [2008], Hopfensitz et Reuben [2009], Wubben *et al.* [2009]) ou bien la dégradation des relations interpersonnelles entre les agents au sein d'un groupe ou d'une entreprise.

(ii) Les mécanismes de punitions engendrés par ces émotions représentent par ailleurs un coût économique non négligeable pour ceux qui les mettent en œuvre. Il en résulte qu'en termes de bien-être social la mise en œuvre d'une procédure de sanction n'a pas toujours l'effet escompté même lorsqu'elle rectifie les comportements opportunistes. Morel [2012] identifie par ailleurs les effets délétères de la punition au sein des organisations lorsque la crainte de la sanction est un frein à la divulgation des erreurs de décisions individuelles.

(iii) Les effets de la sanction morale ne sont en général pas permanents. Dans l'expérience de Fehr et Gächter [2000], le comportement égoïste des joueurs réapparaît massivement et spontanément dès l'élimination de la menace de sanction. Dans le jeu du bien public, un comportement bien identifié de « fin de jeu » – la chute drastique des contributions à l'annonce de la fin du jeu – révèle également cet aspect stratégique. Dans le jeu du dictateur, les effets potentiels de la honte sur la capacité de don sont éliminés lorsque l'on propose aux participants une option de sortie leur permettant de préserver leur intérêt individuel à l'insu du récipiendaire (Dana *et al.* [2006]).

(iv) Par définition, les émotions à valence négative représentent une perte en termes d'utilité pour ceux qui les ressentent. En conséquence, elles nuisent au bien-être individuel. Certains travaux en psychologie soulignent en particulier la corrélation négative entre le sentiment de honte et le bien-être individuel (Tangney *et al.* [2007]).

(v) Enfin, la culpabilité (ou la honte) peuvent également avoir des effets comportementaux induits non souhaités. Une étude récente d'Hooge *et al.* [2011] montre ainsi que la culpabilité a des effets bénéfiques pour la personne à qui l'on a fait du tort, mais représente un désavantage pour les autres personnes de notre environnement social. La réparation du tort (en faveur de la victime) se fait au détriment d'autres personnes (des tiers) : un individu se sentant coupable de ne pas s'être suffisamment occupé d'un de ses proches, peut décider de lui consacrer plus de temps pour compenser son absence. Selon Hooge *et al.* [2011], cependant, ce temps est sélectionné sur du temps réservé à un tiers, et non sur du temps réservé à lui-même.

En dépit de l'impact positif des émotions morales à valence négative dans l'émergence des comportements prosociaux, leurs limites augmentent

l'intérêt que l'on peut porter aux émotions morales dont la valence est positive, et qui ont pourtant été largement moins étudiées dans la littérature.

2.3. La promotion des comportements : la récompense affective

Le nouveau regard porté sur les affects positifs révèle une tendance de la psychologie moderne à sonder (dorénavant) les mécanismes favorables au bien-être de l'individu, ce que l'on appelle la « psychologie positive » (Seligman et Csikszentmihályi [2000]). Les émotions morales à valence positive – comme la fierté, la gratitude ou même l'admiration – possèdent cette propriété, que l'on retrouve également dans les travaux sur « l'économie du bonheur » (Davoine [2012]). Elles sont, par définition, une source de bien-être pour ceux qui les ressentent. Elles peuvent induire des mécanismes de récompenses qui favorisent le respect, l'adoption et la diffusion des normes sociales. Elles sont également susceptibles de jouer sur la nature des relations interpersonnelles au sein des institutions.

2.3.1. Le rôle ambigu de la fierté

Du point de vue de la théorie économique standard, être fier de l'efficacité de son travail sans demander une contrepartie financière, renoncer à bénéficier de dons financiers en provenance d'œuvres de charité, ou rejeter des offres profitables mais perçues comme « offensantes » (Fehr et Schmidt [2005]), traduit *de facto* une forme d'irrationalité. Dans le cadre d'une négociation ou d'un processus de coopération, cependant, la fierté favorise l'établissement d'une réputation positive qui promeut de meilleurs résultats sur le long terme (Lea et Webley [1997]). Andreoni et Petrie [2004] montrent ainsi, dans un jeu du bien public, que la fierté accentue le « goût de donner » et renforce en conséquence l'attitude positive de ceux qui contribuent davantage que les autres. Dans un contexte réel, ceci a une incidence directe sur les comportements d'économies d'énergie lorsque l'individu peut étalonner son comportement sur celui de ses proches voisins (Schultz *et al.* [2007]). L'efficacité de la fierté semble ainsi être conditionnée à la visibilité de l'acte par autrui. Liée étroitement à la notion d'accomplissement et à l'estime de soi, la fierté a également un rôle central dans la relation au travail : elle est une source de motivation intrinsèque (Bénabou et Tirole [2006]), elle favorise l'accomplissement (l'effort) de l'individu dans le travail, et est associée à une forte capacité d'autonomie au sein des organisations.

2.3.2. La chaîne vertueuse de la gratitude

La gratitude est une émotion que l'on ressent lorsqu'une autre personne a effectué (ou a tenté de le faire) un acte, de manière intentionnelle et volon-

taire, qui nous procure un gain précieux (un don d'argent, de l'aide, etc.). On évoque la gratitude lorsque l'acte représente un coût pour la personne qui le réalise, sans qu'il y ait de compensation financière (immédiate) de la part de celui qui le reçoit (McCullough *et al.* [2001]). A la différence d'un simple sentiment de joie, la gratitude renforce les comportements prosociaux lors d'interactions économiques (Barlett et DeSteno [2006]; De Steno *et al.* [2010]). La gratitude est en ce sens associée à la notion plus répandue en économie de réciprocité (voir Bonein [2008] pour une revue).

Dans un jeu similaire au dilemme du prisonnier, l'induction de la gratitude – un complice de l'expérimentateur prodiguant gratuitement une aide au sujet – augmente significativement le taux de coopération des sujets (De Steno *et al.* [2010]). L'étude de Grant et Gino [2010] confirme ces résultats en montrant également que l'expression de la gratitude envers « l'aidant » pérennise le comportement prosocial de ce dernier. L'expression concrète de la gratitude, de l'aidé à l'aidant, sous la forme d'un simple remerciement ou d'une réciprocité positive, valorise socialement les aidants, initie un sentiment de fierté (2.3.1), et renforce leur désir de continuer à adopter des comportements utiles pour la société. La gratitude, source de bien-être pour celui qui la ressent, augmente ainsi la probabilité que la personne poursuive son comportement altruiste et bénévole (Konow et Earley [2008]).

L'effet bénéfique de la gratitude dépasse cependant celui de comportements purement réciroques. Dans l'étude de De Steno *et al.* [2010], l'individu témoigne sa reconnaissance aussi bien lorsque le jeu s'effectue avec la personne (le complice) qui l'a aidé précédemment que lorsqu'il s'agit d'une autre personne, inconnue, qui n'a pas de rapport avec l'aide obtenue par le sujet. Il semble donc que la gratitude puisse provoquer un effet « d'entraînement » positif en incitant les individus à se comporter également de manière prosociale avec ceux qui ne sont pas responsables de l'émergence de cette émotion. La gratitude a donc un effet direct de réciprocité positive mais aussi un effet indirect tourné vers une personne tiers qui n'est pas responsable du sentiment de gratitude. Ce comportement de réciprocité indirect révèle la capacité de la gratitude à engendrer une « chaîne » de comportements prosociaux (altruisme, coopération, confiance, etc.) qui sont bénéfiques pour les relations interpersonnelles entre les agents au sein de l'entreprise ou des institutions.

2.3.3. L'admiration, une émotion à distance

L'émotion d'admiration peut susciter des comportements similaires à ceux observés à la suite de l'expression de la gratitude. L'étude de Schnall *et al.* [2010] montre ainsi que les individus ayant subi l'induction du sentiment d'admiration sont, comparativement à ceux qui sont dans un état émotionnel neutre, plus susceptibles de participer volontairement à une étude non payée ou à aider l'expérimentateur bénévolement à réaliser une tâche ardue. L'admiration a, selon Haidt [2003], un caractère plus général qui dépasse le cadre d'un rapport bi ou tri latéral. Ainsi, alors que l'émergence de la gratitude exige que l'on soit directement affecté par le geste moral adopté par un tiers, l'admiration peut naître du simple fait que l'on a été

témoin de cet acte. En ce sens, l'effet prosocial de l'admiration est potentiellement plus ample que celui de la gratitude qui est davantage une émotion de proximité. En revanche, les conditions d'émergence de l'admiration sont sans doute plus difficiles à réaliser que celles de la gratitude.

2.3.4. Le processus empathique

Le processus d'empathie correspond, en psychologie morale, au processus conduisant à voir les événements de l'intérieur (Batson et Shaw [1991]). Davantage qu'une émotion purement discrète, l'empathie est un processus affectif à travers lequel les individus ressentent et partagent les émotions qu'ils perçoivent chez les autres (Kedia [2009]). Comme le souligne Sugden [2002], l'empathie est proche du concept de « sympathie », fondation de la moralité chez les penseurs moralistes du 18^e siècle (Adam Smith [1759] et David Hume [1739]). En psychologie, la sympathie correspond à une réponse émotionnelle déclenchée par l'émergence d'une émotion chez l'autre, sans que cette réponse soit nécessairement identique à l'émotion de l'autre (contrairement à ce qui se passe avec l'empathie).

Ces deux processus, empathie et sympathie, sont déclenchés par la perception de la souffrance ou de la tristesse chez l'autre, qu'il soit proche ou étranger. Elles conduisent généralement les personnes qui les ressentent à aider, à réconforter et à tenter de diminuer la souffrance d'autrui. Ceci leur confère donc un rôle majeur vis-à-vis de l'adoption de comportements prosociaux ou moraux. Des individus possédant une personnalité empathique ont ainsi tendance à adopter des conduites prosociales – participation à des œuvres de charité, aide à des personnes en difficulté (Kedia [2009]). Inversement, les psychopathes, souffrant d'un déficit empathique, sont incapables d'anticiper les émotions de condamnation d'autrui (colère ou dégoût), ne perçoivent pas le tort qu'ils commettent vis-à-vis des autres et n'en ressentent, en conséquence, aucune forme de culpabilité. L'empathie serait ainsi fortement corrélée au sentiment de culpabilité (Tangney *et al.* [2007]). L'empathie joue enfin un rôle moteur dans l'émergence de la gratitude (2.3.2). Celle-ci nécessite en effet de pouvoir évaluer l'intention sous-jacente au geste réalisé par le bienfaiteur, aptitude qui transite par le processus d'empathie (McCullough *et al.* [2001]). Le bienfaiteur peut également être motivé par l'anticipation empathique de la gratitude chez le bénéficiaire et par le plaisir qu'il ressentira en accomplissant un acte prosocial.

En économie, le rôle de l'empathie dans l'émergence des comportements prosociaux est l'objet d'une littérature naissante³. Andreoni et Rao [2011] proposent une expérimentation dans laquelle les sujets de la condition « Empathie » doivent imaginer leurs réactions et leurs choix pour chacun des différents rôles d'un jeu du dictateur, avant que les rôles définitifs ne soient déterminés par le hasard. Les dictateurs qui prennent virtuellement la place du récipiendaire sont ainsi beaucoup plus généreux que ceux qui n'ont

3. Voir notamment Stark et Falk [1998], Singer et Fehr [2002], Sally [2001], Stahl et Haruvy [2006], Manner et Gowdy [2010], Andreoni et Rao [2011], Zak [2011a, 2011b], Molnar-Szakacs [2011], Calvet et Alm [2014].

pas cette possibilité. Calvet et Alm [2014] étudient le rôle de l'empathie dans le respect des obligations fiscales. Les auteurs proposent un modèle dans lequel l'empathie et la sympathie jouent sur la force avec laquelle des émotions comme la culpabilité ou l'embarras diminuent le bien-être des individus lorsqu'ils fraudent. Plus une personne est empathique, moins elle est susceptible d'adopter un comportement frauduleux. Cette hypothèse est testée dans un jeu de déclaration d'impôts qui suit une procédure d'induction de l'empathie. Dans la condition « Empathie », les participants doivent écrire leur propre définition d'une règle morale mettant l'accent sur l'empathie (« Ne fais pas aux autres ce que tu ne voudrais pas que l'on te fasse ») après avoir lu plusieurs versions de l'expression de cette règle d'or dans différentes religions. Les résultats indiquent que les sujets de la condition « Empathie » succombent moins fréquemment à la tentation de tricherie (qui consiste à déclarer moins de revenu pour diminuer leur niveau d'imposition) que les sujets de la condition « Contrôle ».

2.4. Les enseignements principaux

Une abondante littérature en économie expérimentale s'interroge sur les conditions d'existence et d'émergence de comportements prosociaux ou moraux. De nombreuses preuves expérimentales indiquent que les individus adoptent à la fois des comportements égoïstes (dans certaines situations) mais également des comportements altruistes, honnêtes, coopératifs, équitables, etc. Ces comportements sont étroitement liés à l'émission d'émotions morales et dépendent du contexte de la décision, des croyances des individus relatives aux intentions ou aux attentes d'autrui ainsi que de facteurs de personnalité. Pour la théorie économique, la prise en compte des émotions morales représente par conséquent un défi au regard du cadre standard de la rationalité économique. Nous regardons comment les travaux de modélisation récents tentent de relever ce défi théorique.

3. La prise en compte des émotions morales dans la théorie économique

C'est essentiellement dans le cadre des modèles de « préférences sociales » (Camerer [2003], Serra [2007]) que les émotions morales ont été introduites dans l'analyse économique. L'objectif de ces nouveaux modèles est de rendre compte des comportements observés en laboratoire tout en respectant la rationalité économique et le principe de l'individualisme méthodologique. Dans cette optique, deux façons d'intégrer les émotions morales dans la théorie standard sont envisageables. On peut tout d'abord s'appuyer sur la caractéristique la plus visible des émotions, à savoir leur valence, en intégrant les affects sous la forme de paramètres exogènes dans la fonction

d'utilité individuelle traditionnelle. Les émotions sont dans ce cas considérées comme des préférences exogènes (Guala et Mittone [2010]) et sont activées parce que préexistent implicitement des normes sociales et morales qui contraignent les choix des individus (3.1). En prenant appui sur une conception élargie de ce que recouvre l'émotion, d'autres types de modélisations introduisent l'idée que l'influence des émotions sur les comportements individuels dépend, comme le souligne la littérature expérimentale, des intentions des agents, de leurs attentes et de leur personnalité, du contexte dans lequel s'effectue le processus de décision et, enfin, de la présence plus explicite de la norme sociale. Au lieu de considérer la préférence morale comme une donnée (via un paramétrage exogène de l'émotion), ces modélisations considèrent *a contrario* les émotions morales comme des préférences endogènes (3.2).

3.1. L'intégration des émotions morales comme préférences exogènes

Dans un modèle économique standard, chaque individu maximise sa propre fonction d'utilité qui dépend directement de ses gains monétaires. Dans de nombreux modèles comportementaux, comme ceux d'Andreoni [1990], de Fehr et Schmidt [1999] ou de Bolton et Ockenfels [2000], l'introduction d'une préférence sociale ou morale est effectuée en intégrant, dans la fonction d'utilité des agents, un ou plusieurs paramètres exogènes prenant en compte le bien-être ou les gains d'autrui. Dans le modèle séminal d'Andreoni [1990], cette préférence correspond à un « goût de donner », que l'on peut associer à un affect positif et qui mesure le plaisir qu'un individu retire lorsqu'il choisit de participer à la production d'un bien public (construction d'une école, œuvre caritative, participation à un projet collectif, etc.). Dans les articles fondateurs de Fehr et Schmidt [1999] et Bolton et Ockenfels [2000], un arbitrage équivalent est effectué au sein de la fonction d'utilité, entre un plaisir égoïste et une contrainte exogène associée à une « préférence pour l'équité ».

Ce type de modèles comportementaux a connu un fort engouement en raison de leur capacité à rendre compte de l'existence de comportements individuels observés en laboratoire à partir des outils standards de l'économiste. Les émotions morales sont mobilisées pour donner une explication « plausible » des préférences des agents – la gratitude pour le « goût de donner », l'envie, la colère, la culpabilité, la gêne ou encore la honte pour « l'aversion à l'iniquité » – sans que les fondements émotionnels de ces préférences ne soient pour autant précisés. Dans la mesure où les modèles ne tiennent compte que de la valence de l'émotion (le caractère désagréable ou agréable associé à la conséquence de la décision) et qu'ils n'intègrent aucune interaction sociale effective entre les individus (au-delà des gains), ils ne permettent pas de comprendre comment le processus émotionnel conduit au processus de choix. S'agissant d'une émotion morale, il faudrait notamment savoir pourquoi l'individu considère-t-il la règle morale implicite (l'équité ou le goût de donner) comme incontournable (invariante) dans

toutes les situations ? A quelle condition sa violation (ou son respect) conduit-elle l'individu à nourrir des émotions négatives (ou positives) ? Pour quelle raison le ressenti de ces émotions ou leur anticipation modifient-ils la « tendance à l'action » et le choix opéré par l'agent économique ? Dans quelle mesure le comportement moral dépend-il du contexte de la décision, de la présence d'anonymat ou encore du caractère observable des choix ? Quelle relation sociale ou morale existe-t-il, au-delà de l'espace des gains, entre les individus qui sont à l'initiative des processus de choix ?

Les modèles « d'équité distributionnelle » ou ceux associés au « goût de donner » présentent un caractère *ad hoc* (Wolfesperger [2001]) au sens où ils expliquent le comportement équitable ou altruiste en posant de façon exogène l'existence d'une « préférence pour l'équité » ou un « plaisir de donner ». On pourrait ainsi décliner la moralité des individus du point de vue formel en postulant, de façon équivalente, un goût pour « l'honnêteté », pour la « vertu » ou la « sincérité » ; ou encore, par opposition, supposer l'existence d'un « plaisir à la malveillance » (malice) ou à la « dissimulation ». Cette limite, le caractère *ad hoc* de la formalisation qui ne permet pas d'apprécier le fondement de la motivation évoquée (le sens de l'équité), découle directement de la façon dont sont introduites les émotions morales. Si l'on ne tient pas compte notamment du fait que l'émergence des émotions résulte de l'intériorisation par les individus de règles de comportement, on sous-estime le fait que ces règles sont davantage des préférences endogènes que des contraintes exogènes qui s'imposent à l'individu (Stringham [2011]). Dans ce cas, la représentation des préférences morales doit reposer sur une base contextuelle dans laquelle les préférences dépendent de façon cruciale de l'interaction sociale entre l'individu et autrui et du contexte dans lequel s'effectue la prise de décision. C'est précisément ce que proposent d'autres types de modèles comportementaux basés notamment sur le rôle des intentions des acteurs ou de leurs attentes.

3.2. L'intégration des émotions morales comme préférences endogènes

Une autre façon, plus réaliste, de modéliser les comportements prosociaux est de considérer que les individus adoptent des règles de comportement endogènes. Ces modèles comportementaux s'appuient sur une littérature économique expérimentale qui a montré, à la suite de Blount [1995], le rôle explicatif fondamental des intentions dans la prise de décision. Dans cette optique, un individu accepte plus volontiers un partage non équitable s'il sait que la décision n'est pas le fruit d'un acte intentionnel (Falk *et al.* [2003], Sutter [2007]). Dès lors, il importe de prendre en compte l'intention qui préside à l'action dans l'analyse du comportement prosocial (3.2.1). Au-delà de l'intentionnalité, d'autres auteurs soulignent également le rôle fondamental que jouent les attentes d'un individu – mon partenaire escompte-t-il (ou non) que j'adopte un comportement égoïste ou prosocial ? – auxquelles s'associent le contexte de la décision ainsi que la personnalité des acteurs (3.2.2). Enfin, l'endogénéisation des émotions morales modifie

également la façon dont la norme sociale évolue au sein d'une collectivité d'individus. En modélisant la norme sociale de façon plus explicite que dans les modèles précédents, il devient possible d'analyser comment les émotions morales et la norme évoluent au sein de la collectivité (3.2.3).

3.2.1. Le rôle fondateur de l'intention

Le modèle séminal le plus novateur est celui de « l'équité intentionnelle » de Matthew Rabin [1993]⁴. Afin de formaliser cette conception de l'équité intentionnelle, Rabin [1993] modifie la théorie des jeux standard en faisant dépendre la fonction de gains des joueurs, non seulement de leurs actions, mais de leurs croyances. Le modèle s'inspire des premiers « jeux psychologiques » (Geanakopulos *et al.* [1989]) dans lesquels l'utilité des joueurs dépend en partie de déterminants affectifs. Au centre du modèle se trouve l'idée de réciprocité positive ou négative qui dépend des intentions bien ou malveillantes de son partenaire. Pour en saisir l'intuition, plaçons-nous, comme le fait Meidinger [2000], dans le cas le plus simple du dilemme du prisonnier. Supposons que les gains de chaque joueur soient égaux à 4 lorsqu'ils coopèrent tous les deux et à 1 s'ils font défection tous les deux. Lorsque, enfin, l'un des deux coopère tandis que l'autre fait simultanément défection, ce dernier gagne 6 tandis que celui qui coopère ne gagne rien. En théorie des jeux, un joueur rationnel a intérêt à faire défection quelle que soit la stratégie qu'il envisage pour l'autre joueur, l'équilibre en stratégies dominantes étant donc la défection des deux joueurs. Tout peut changer si l'on introduit le rôle des intentions et des croyances. Admettons que j'imagine être face à un joueur bienveillant. J'anticipe de sa part une attitude coopérative. Si je ne maximise que mon intérêt matériel, j'ai toujours intérêt à faire défection. Si, en revanche, je prends en compte, dans ma fonction d'utilité, le sentiment endogène positif de sympathie associé à l'attitude (anticipée) de mon partenaire, ceci me conduit par réflexivité à porter un jugement moral sur les motifs de mes propres actions. Dès lors, il devient moralement coûteux de faire défection en face d'un joueur bienveillant et *a contrario* moralement bénéfique de coopérer avec celui-ci. Si les deux joueurs sont dans le même état d'esprit, un équilibre du jeu de Nash psychologique, en termes de croyances et d'actions, conduit à un équilibre de bienveillance réciproque dans lequel la coopération s'instaure. En revanche, dans le cas contraire, face à un individu que je crois malveillant (qui jouerait la défection), le raisonnement s'inverse et les considérations émotionnelles motivent un autre équilibre réciproque porté vers la défection simultanée. Le jeu du dilemme du prisonnier se transforme dans ce cas en un jeu de coordination (analogue au jeu de la « chasse au cerf ») possédant deux équilibres de Nash dont l'un, celui de coopération réciproque, domine strictement l'autre, celui de défection réciproque. On peut alors envisager, comme le propose Meidinger [2000, p. 47], une « solution humienne » qui impliquerait qu'une communication crédible d'intention, une promesse, s'instaure entre les deux

4. Voir Dufwenberg et Kirchsteiger [2004] et Falk et Fischbacher [2006] pour des extensions récentes de ce modèle.

joueurs. Les résultats expérimentaux indiquent en effet que les joueurs qui anticipent que les autres vont coopérer (ou faire défection) ont tendance à coopérer (à faire défection) davantage eux-mêmes. Ici, les émotions morales favorisent une communication verbale ou non verbale qui permet au joueur de signaler, via ses expressions émotionnelles, la nature de ses intentions. L'identification, le sourire ou la communication améliorent ainsi la coopération dans le jeu du bien public en rendant plus transparentes les intentions des joueurs (Charness et Dufwenberg [2006]).

3.2.2. Attentes, contexte et personnalité

L'apport central du modèle de Rabin [1993] est d'avoir mis en évidence le caractère conditionnel de la prise de décision souligné par les différents résultats expérimentaux issus des jeux de négociation ou de coopération : dans tel type d'environnement, je choisis l'action C si « l'évènement » V se réalise. J'initie, par exemple, une coopération initiale si et seulement si j'anticipe que mon partenaire fera de même. L'intention joue ici un rôle crucial, mais pas seulement. L'ensemble V peut en effet désigner l'espace des actions anticipées mais, tout aussi bien, celui des attentes des joueurs, ou encore celui associé au contexte de la décision.

Le rôle des attentes est crucial notamment dans la relation de négociation lorsque le pouvoir de décision entre les acteurs est asymétrique. Par exemple, dans un jeu où l'un des partenaires n'a aucun pouvoir de veto (et donc aucune décision à prendre), son opposant devrait être entièrement libre de toute contrainte morale. Les nombreux résultats expérimentaux du « jeu du dictateur » montrent cependant qu'il n'en est rien (Petit et Rouillon [2010]). Le fait que l'individu raisonne aussi par rapport aux attentes d'autrui et qu'il soit sensible au contexte de la décision permet d'expliquer en partie ce décalage. L'autre joueur escompte-t-il une décision équitable ? Ma décision sera-t-elle connue publiquement ?

Les résultats expérimentaux ont donc conduit à intégrer les attentes des agents économiques dans les modèles comportementaux. Rotemberg [2008] modélise ainsi, dans un jeu de l'ultimatum, un seuil minimal d'altruisme attendu de son partenaire en dessous duquel tout comportement excessivement opportuniste déclenche la réaction de rejet du joueur qui en subit les conséquences. Dans le même esprit, Battigalli et Dufwenberg [2007, 2009] ont construit, à la suite de Charness et Dufwenberg [2006], un modèle prometteur dans lequel « l'aversion à la culpabilité » (« *guilt aversion* ») naît du décalage entre ce que je pense que l'autre attend et ce que je suis prêt à lui concéder. Concrètement, l'individu anticipe la déception de son partenaire lorsqu'il croit que ce dernier obtient moins (au cours de l'échange) que ce qu'il pouvait espérer. Il en résulte un sentiment de culpabilité dont l'amplitude dépend de la sensibilité de son émetteur ainsi que de la « personnalité » de son partenaire (s'agit-il d'un inconnu ou d'un proche ?).

Le rôle de la sensibilité affective se retrouve également dans le modèle, moins cité dans la littérature, de Bowles et Gintis [2005]. Dans leur modèle illustratif du jeu du bien public, la culpabilité naît de la confrontation entre le niveau de contribution du joueur et le niveau de contribution qu'il considère

comme étant la norme sociale. La culpabilité, en raison de son caractère privé, est donc ici indépendante (de l'anticipation) des attentes de l'autre joueur mais dépend tout de même d'une attente sociale intériorisée par l'individu. Le modèle tient compte également du sentiment de honte, issu de la désapprobation des autres contributeurs qui sanctionnent un comportement jugé opportuniste. La honte, du fait de son caractère public, et *a contrario* du sentiment de culpabilité, provient d'une interaction avec autrui. Une action punitive ne provoque cependant pas systématiquement la honte de celui qui la subit puisque l'on suppose que chaque individu a une perception différente de ce qu'est la norme de comportement adaptée à la situation de jeu : une punition conduit ainsi au ressenti d'une rancune lorsque le joueur considère sa propre contribution comme suffisamment élevée.

Tadelis [2008], Neilson [2009], Dessi et Zhao [2010] renforcent le poids accordé, dans le processus de décision, au contexte et à la sensibilité affective. Dans le modèle de théorie des jeux psychologique proposé notamment par Tadelis [2008], l'individu dispose d'une sensibilité à ressentir de la honte qui dépend de son type ou de sa personnalité. Le sentiment de honte est également en lien avec le niveau d'exposition des actions choisies par les joueurs. Dans le jeu de la confiance (voir l'arbre du jeu ci-dessous), si un individu a la possibilité, à la première étape du jeu, de me transférer un certain montant d'argent, mon choix (en tant que joueur 2) de lui renvoyer (ou non) une partie de cette somme dépend de la visibilité de mon acte. Si je suis sensible à son approbation morale, et que ma décision (C ou D) est connue avec certitude par le joueur 1, le sentiment de honte influence mon comportement social dans le sens de la réciprocité. Dans le même temps, ceci a également un impact positif sur la confiance que le joueur 1 m'accorde.

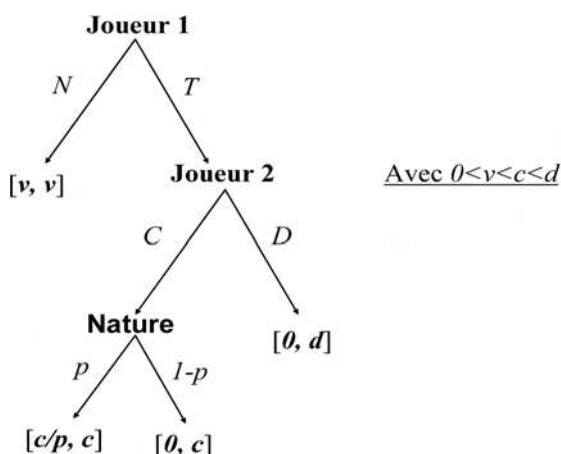


Figure 1. Arbre du jeu de confiance dans le modèle de Tadelis [2008]

Inversement, si le joueur 1 n'a pas connaissance du caractère volontaire de mon acte – il n'a rien obtenu mais il ne sait pas si c'est le fait de ma décision opportuniste (D) ou simplement la conséquence du hasard (C avec la probabilité $1-p$) – j'aurai tendance à croire qu'il peut tout aussi bien me prêter de bonnes intentions quel que soit mon choix. Dans ma fonction d'utilité, l'absence de visibilité minore donc le poids que j'accorde au fait que le joueur 1 attribue une probabilité positive associée à mon comportement opportuniste (D). En conséquence, le faible niveau d'exposition renforce ma possibilité d'obtenir le gain maximal (d) sans souffrir pour autant de la désapprobation morale. L'efficacité de la honte dépend ainsi du niveau d'exposition de l'acte et de mon aversion personnelle à la honte.

3.2.3. La prise en compte explicite de la norme sociale

En intégrant les intentions, les attentes, le contexte ou les facteurs de personnalité, les modèles précédents (3.2.1 et 3.2.2) prolongent l'analyse rationnelle en soulignant, au sein de la fonction d'utilité individuelle, l'existence d'un arbitrage entre un comportement égoïste et une attitude prosociale. Les préférences sociales dépendent des émotions morales de façon endogène : le comportement individuel provient de la désapprobation d'autrui, qui induit un sentiment de honte, ou d'une norme sociale intériorisée, qui conduit à de la culpabilité. Ces modèles intègrent des mécanismes psychologiques précis qui permettent de différencier, au niveau de l'individu, comment une norme sociale implicite module son comportement. En revanche, l'absence de modélisation explicite de la norme ou de la règle sociale qui s'impose au niveau de la collectivité implique que l'émotion morale semble posséder une dimension entièrement subjective spécifique à l'agent. L'objectif des modèles présentés ci-dessous est précisément de rendre explicite la norme sociale et, également, de considérer que son influence résulte de la confrontation réelle des individus avec autrui.

Les premiers modèles mettant l'accent sur la spécification d'une norme bien identifiée sont issus des travaux novateurs d'Akerlof [1980], de Kandell et Lazear [1992] ou de Holländer [1990]. L'intérêt du modèle d'Akerlof [1980] est qu'il rend compte de deux types de motivation : (1) une motivation instrumentale qui incite les individus à respecter la norme en cours dans un certain but (par exemple, obtenir la réciprocité ou conforter sa réputation) et (2) une motivation intrinsèque induite par une disposition à respecter la norme sans attente d'une contrepartie. Dans ce cadre, la désapprobation morale augmente le coût psychologique et affectif associé à la violation de la norme (culpabilité) et accentue également la pression sociale qui renforce le calcul stratégique de l'agent (*i.e.* la préservation de sa réputation). En dynamique, le maintien de la norme (ou au contraire sa disparition) dépend crucialement du nombre de ceux qui l'ont intériorisée (et qui donc croient à son bien-fondé) relativement à ceux qui la respectent uniquement par souci stratégique. Ainsi, entre deux périodes, si, initialement, le nombre d'agents qui respectent la norme est supérieur au nombre de ceux qui l'ont intériorisée, à la période suivante, le nombre d'agents qui croient à son bien-fondé augmente. Or, initialement, le nombre d'agents qui respectent la norme

dépend de l'intensité avec laquelle les agents qui l'ont intériorisée souffrent de ne pas la respecter, *i.e.* de l'intensité de la culpabilité consécutive à la violation de la norme. Ce nombre dépend également, pour chaque agent (qu'ils croient ou non dans la norme) de l'intensité de la perte en termes de réputation consécutive au fait de ne pas respecter la norme. Ce coût en termes de réputation est proportionnel au nombre d'agents qui croient en son bien-fondé et on peut interpréter cela comme étant l'expression d'émotions morales négatives chez ces individus « de principe » qui considèrent alors de manière négative (colère, mépris, dégoût) ceux qui ne respectent pas la norme, ce qui affecte leur réputation. Dans ce modèle, il apparaît donc un lien dynamique entre l'intensité des émotions morales ressenties et exprimées dans un groupe d'agents, et l'influence, le maintien et même la persistance de la norme de comportement qui est initialement présente au sein de ce groupe. Le modèle d'Akerlof [1980] illustre ainsi plutôt bien la dynamique des affects et des habitudes de comportement qui sont à l'origine de l'évolution d'une norme (que l'on pense, par exemple, en France, au débat sur la peine de mort ou, plus récemment, sur la parité).

On retrouve l'articulation entre les émotions morales et la norme dans le modèle d'Holländer [1990] appliqué à un jeu du bien public. La norme de comportement standard émerge à la suite des interactions sociales et du positionnement moral de chacun face à la contribution d'autrui. Le sentiment d'approbation naît en particulier lorsque le niveau de contribution observé pour un individu est supérieur à celui de la contribution moyenne. Davantage, ce sentiment est plus fort lorsque l'écart entre les deux niveaux de contribution est élevé, ce qui traduit un effet de comparaison récemment identifié dans le cas de l'émotion de gratitude (Wood, Brown et Maltby [2011]). Holländer [1990] montre l'existence d'un équilibre « d'échange social » – comparable au mécanisme de la main invisible – dans lequel les contributions individuelles correspondent aux contributions moyennes et le taux d'approbation individuel est cohérent avec les comportements observés. Le modèle récent proposé par Lopez-Pérez [2008] présente quelques similitudes avec ceux d'Akerlof [1980] et d'Holländer [1990]. Dans son approche, coexistent au sein de la société des individus opportunistes et des individus « de principe » qui sont sensibles au respect des normes sociales morales. La norme revêt, dans ce modèle, un statut générique (valable pour toutes les normes) qui tient compte du sens de l'équité des individus mais également de leur attention à l'efficacité économique des choix. Le respect de la norme dépend, pour les agents « de principe », d'un arbitrage entre gains matériels et coût affectif associé à sa violation. Ce coût affectif, lié aux émotions de culpabilité et de honte, dépend de l'intensité avec laquelle l'individu a internalisé la norme, formalisée sous la forme d'un coefficient constant, mais également, de manière assez originale, du nombre de transgresseurs au sein de la population. Posé simplement, plus nombreux seront les agents qui respectent la norme dans l'économie, plus la désutilité provenant de la honte ressentie par l'agent lorsqu'il viole la norme sera importante. On retrouve ici, une possibilité innovante de rendre l'adhésion à la norme conditionnelle, en fonction de paramètres de personnalité mais aussi en fonction de la situation, au sens de la constitution collective de l'économie, dans laquelle le processus de décision s'effectue.

3.3. Le tableau général qui en ressort

A la suite des preuves expérimentales identifiées en laboratoire, de nombreux travaux de modélisation rendent compte du rôle des émotions morales dans l'émergence des comportements prosociaux. Dans cette littérature, les émotions sont intégrées en tant que préférences exogènes, sous la forme de paramètres, ou en tant que préférences endogènes, lorsque leur émergence dépend des intentions, des attentes ou du contexte de la décision.

Dans l'ensemble, peu d'analyses introduisent la norme sociale de façon explicite. Dans les modélisations exogènes de l'émotion, la norme est implicite et invariante ; dans les modélisations endogènes, elle est, le plus souvent, intériorisée. Est donc laissée en suspens la question posée par Elster [1999] du rôle des émotions morales dans l'édification des normes sociales. Un modèle comportemental plus élaboré devrait donc prendre en compte que les agents suivent des normes sociales ou morales qui sont elles-mêmes le produit de processus affectifs. Ceci implique cependant de définir plus précisément la façon dont l'individu interagit avec autrui ou avec le groupe en réponse à la norme.

4. Conclusion

Dans cet article, nous avons retracé de façon non exhaustive une littérature abondante traitant du rôle des émotions morales comme la colère, la culpabilité, la honte, mais aussi la gratitude ou la fierté, dans l'émergence des comportements prosociaux. L'intérêt principal de cette littérature est de renouveler la conception standard de l'analyse économique en intégrant, aux côtés de la recherche de l'intérêt individuel, d'autres motivations sociales ou morales comme « l'attention à l'autre » (Robin [2012]), la coopération, la réciprocité, l'altruisme ou la recherche de l'équité.

Les nombreuses preuves expérimentales de l'existence de préférences sociales ont conduit à des travaux de modélisation novateurs. Initialement mobilisées de façon rudimentaire, sous la forme de paramètres exogènes, les émotions morales ont été endogénéisées en tenant compte des apports et de la spécificité de la théorie psychologique des émotions. Ces modèles se caractérisent ainsi par davantage de réalisme par rapport aux mécanismes affectifs internes qui sous-tendent la prise de décision. La complexité du domaine des affects, leur subtilité dans le cas des émotions morales, ainsi que leur implication forte dans le processus de décision, augurent cependant que de nombreux travaux comportementaux investiront à l'avenir le domaine de recherche que nous avons retracé.

Références bibliographiques

- AKERLOF G. A. [1980], A Theory of social custom of which unemployment may be one consequence », *Quarterly Journal of Economics*, 94, p. 749-775.
- ANDREONI J. [1990], Impure Altruism and Donations to Public Goods : a Theory of Warm-Glow Giving, *Economic Journal*, 100, p. 464-477.
- ANDREONI J., PETRIE R. [2004], Public goods experiment without confidentiality : a glimpse into fund-raising, *Journal of Public Economics*, 88, p. 1605-1623.
- ANDREONI J., RAO J. M. [2011], The power of asking : How communication affects selfishness, empathy, and altruism, *Journal of public economics*, 95, p. 513-520.
- BARLETT M. Y., DE STENO D. [2006], Gratitude and Prosocial Behavior, Helping When It Costs You, *Psychological Science*, 17, p. 319-325.
- BATSON C. D., SHAW L. L. [1991], Evidence for altruism : Toward a pluralism of prosocial motives, *Psychological inquiry*, 2, p. 107-222.
- BATTIGALLI P., DUFWENBERG M. [2007], Guilt in Games, *American Economic Review, papers and proceedings*, 97, p. 170-176.
- BATTIGALLI P., DUFWENBERG M. [2009], Dynamic psychological games, *Journal of Economic Theory*, 144, p. 1-35.
- BEN-SHAKHAR G., BORNSTEIN G., HOPFENSITZ A., VAN WIDEN F. [2007], Reciprocity and emotions in bargaining using physiological and self-report measures, *Journal of economic psychology*, 28, p. 314-323.
- BENABOU R., TIROLE J. [2006], Incentives and Prosocial Behavior, *American Economic Review*, 96, p. 1652-1678.
- BLOUNT S. [1995], When social outcomes aren't fair : The effect of causal attributions on preferences, *Organizational Behavior and Human Decision Processes*, 63, p. 131-144.
- BONEIN A. [2008], La réciprocité, entre psychologie et rationalité économique, *Revue française d'économie*, 23, p. 203-232.
- BOLTON G. E., OCKENFELS A. [2000], ERC : A theory of equity, reciprocity and competition, *American Economic Review*, 90, p. 166-193.
- BOSMAN R., VAN WINDEN F. [2002], Emotional hazard in a power-to-take experiment, *Economic journal*, 112, p. 147-169.
- BOWLES S., GINTIS H. [2005], Prosocial Emotions, dans *The Economy as an Evolving Complex System III : Current Perspectives and Future Directions*, BLUME L. E. et DURLAUF S. N. (eds), New York : Oxford University Press, Vol. 3, p. 339-364.
- BROWN B. R. [1970], Face-saving following experimentally induced embarrassment, *Journal of Experimental Social Psychology*, 6, p. 255-271.
- CALVET R., ALM J. [2014], Empathy, sympathy, and tax compliance, *Journal of economic psychology*, 40, p. 62-82.
- CAMERER C. [2003], *Behavioral Game Theory : Experiments in strategic interaction*, Princeton, NJ : Princeton University Press.
- CARPENTER J., DANIERE A., TAKAHASHI L. [2004], Cooperation, Trust, and Social Capital in Southeast Asian Urban Slums, *Journal of Economic Behavior & Organization*, 55, p. 533-551.
- CARPENTER J. P., MATTHEWS P. H. [2012], Norm enforcement : Anger, Indignation or Reciprocity ?, *Journal of the European Economic Association*, 10, p. 555-572.

- CHARNESS G., DUFWENBERG M. [2006], Promises and Partnership, *Econometrica*, 74, p. 1579-1601.
- DANA J., CAIN D., DAWES R. [2006], What you don't know won't hurt me : Costly (but quiet) exit in dictator games, *Organizational Behavior and Human Decision Processes*, 100, p. 193-201.
- DAVOINE L. [2012], *Économie du bonheur*, La Découverte, Repères, Paris.
- DESSI R., ZHAO X. [2010], Self-esteem, shame and personal motivation, *Working paper Series*, n° 639, IDEI, Toulouse School of Economics.
- DE STENO D., BARLETT M. Y., BAUMANN J., WILLIAMS L. A., DICKENS L. [2010], Gratitude as Moral Sentiment : Emotion-Guided Cooperation in Economic Exchange, *Emotion*, 10, p. 289-293.
- DUFWENBERG M., KIRCHSTEIGER G. [2004], A Theory of sequential reciprocity, *Games and Economic Behavior*, 47, p. 268-298.
- ELSTER J. [1998], Emotions and Economic Theory, *Journal of Economic Literature*, 36, p. 47-74.
- ELSTER J. [1999], *Alchemies of the mind : Rationality and the emotions*, Cambridge : Cambridge University Press.
- FALK A., FEHR E., FISHBACHER U. [2003], On the nature of fair behavior, *Economic inquiry*, 41, p. 20-26.
- FALK A., FISHBACHER U. [2006], A theory of reciprocity, *Games economic behavior*, 54, p. 293-315.
- FEHR E., FISCHBACHER U. [2004], Third party punishment and social norms, *Evolution and Human Behavior*, 25, p. 63-87.
- FEHR E., SCHMIDT K. M. [1999], A theory of Fairness, Competition, and Cooperation, *Quarterly Journal of Economics*, 114, p. 817-868.
- FEHR E., GÄCHTER S. [2000], Cooperation and Punishment in Public Good Experiments, *American Economic Review*, 90, p. 980-994.
- FRANK R. H. [1988], *Passions within reason : The Strategic role of emotions*, New York : Norton.
- FRIJDA N. H. [1986], *The Emotions*, Cambridge, Cambridge University Press.
- GRANT A. M., GINO F. [2010], A little Thanks Goes a Long Way : Explaining Why Gratitude Expressions Motivate Prosocial Behavior, *Journal of personality and social psychology*, 98, p. 946-955.
- GEANAKOPOLOS J., PEARCE D., STACCHETTI E. [1989], Psychological games and sequential rationality, *Games and Economic Behavior*, 1, p. 60-79.
- GIORGETTA C., GRECUCCI A., BONINI N., CORICELLI G., DEMARCHI G., BRAUN C., SANFEY A. G. [2013], Waves of regret : A meg study of emotion and decision-making, *Neuropsychologia*, 51, p. 38-51.
- GUALA F., MITTONE L. [2010], Paradigmatic experiments : The Dictator Game, *Journal of Socio-Economics*, 39, p. 578-584.
- HAIDT J. [2001], The emotional dog and its rational tail : A social intuitionist approach to moral judgment, *Psychological Review*, 108, p. 814-834.
- HAIDT J. [2003], The Moral Emotions, in Davidson R. J., Scherer K.R., Goldsmith H. H. (Eds.), *Handbook of affective sciences*, Oxford University Press.
- HOLLÄNDER H. [1990], A Social Exchange Approach to Voluntary Cooperation, *American Economic Review*, 80, p. 1157-1167.

- HOOGE I. E., ZEELENBERG M., BREUGELMANS S. M. [2007], Moral sentiments and cooperation : Differential influences of shame and guilt, *Cognition and emotion*, 21, p. 1025-1042.
- HOOGE I. E., BREUGELMANS S. M., ZEELENBERG M. [2008], Not So Ugly After All : When Shame Acts as a Commitment Device, *Journal of personality and social psychology*, 95, p. 933-943.
- HOOGE I. E., NELISSEN R. M. A., BREUGELMANS S. M., ZEELENBERG M. [2011], What Is Moral About Guilt ? Acting « Prosocially » at the Disadvantage of Others, *Journal of personality and social psychology*, 100, p. 462-473.
- HOPFENSITZ A., REUBEN E. [2009], The importance of emotions for the effectiveness of social punishment, *The Economic Journal*, 119, p. 1534-1559.
- HUME D. [1739], 1991, *Le Traité de la Nature Humaine*, Tome 2, Les Passions, Flammarion.
- KANDEL E., LAZEAR E. P. [1992], Peer Pressure and Partnership, *Journal of Political Economy*, 100, p. 801-817.
- KEDIA G. [2009], La morale et les émotions, *Revue électronique de psychologie sociale*, 4, p. 47-53.
- KETELAAR T., AU W. T. [2003], The effects of feelings of guilt on the behavior of uncooperative individuals in repeated social bargaining games : An affect-as-information interpretation of the role of emotion in social interaction, *Cognition and emotion*, 17, p. 429-453.
- KIRCHSTEIGER G., RIGOTTI L., RUSTICHINI A. [2006], Your morals might be your moods, *Journal of Economic Behavior and Organization*, 59, p. 155-172.
- KONOW J., EARLEY J. [2008], The hedonistic paradox : is *homo economicus* happier ?, *Journal of Public Economics*, 92, p. 1-33.
- LEA S. G. E., WEBLEY P. [1997], Pride in economic psychology, *Journal of Economic Psychology*, 18, p. 323-340.
- LOPEZ-PÉREZ R. [2008], Aversion to norm-breaking : a model, *Games and economic behavior*, 64, p. 237-267.
- MANNER M., GOWDY J. [2010], The evolution of social and moral behavior : Evolutionary insights for public policy, *Ecological economics*, 69, p. 753-761.
- MASCLET D., NOUSSAIR C., TUCKER S., VILLEVAL M.-C. [2003], Monetary and Non-monetary Punishment in the Voluntary Contributions Mechanism, *American Economic Review*, 93, p. 366-380.
- MCCULLOUGH M. E., KILPATRICK S. D., EMMONS R. A., LARSON D. B. [2001], Is Gratitude a Moral Affect ?, *Psychology Bulletin*, 127, p. 249-266.
- MEIDINGER C. [2000], Vertus artificielles et règles de justice chez Hume : une solution au dilemme du prisonnier en termes de sentiments moraux, *Revue de Philosophie Economique*, 1, p. 33-50.
- MOLNAR-SZAKACS I. [2011], From actions to empathy and morality – A neural perspective, *Journal of Economic Behavior & Organization*, 77, p. 76-85.
- MOREL C. [2012], *Les décisions absurdes II – Comment les éviter*, Bibliothèque des Sciences Humaines, Gallimard.
- MORETTI L., DI PELLEGRINO G. [2010], Disgust Selectively Modulates Reciprocal Fairness in Economic Interactions, *Emotion*, 10, p. 169-180.
- NEILSON W. [2009], A theory of kindness, reluctance, and shame for social preferences, *Games and Economic Behavior*, 66, p. 394-403.

- PEETERS R., VORSATZ M. [2013], Immaterial rewards and sanctions in a voluntary contribution experiment, *Economic Inquiry*, 51, p. 1442-1456.
- PETIT E. [2009], Le rôle des affects en économie, *Revue d'Économie Politique*, 119, p. 859-897.
- PETIT E. [2010], La négociation : Les enseignements du jeu de l'ultimatum, *Négociations*, 13, p. 77-96.
- PETIT E., ROUILLON S. [2010], La négociation : Les enseignements du jeu du dictateur, *Négociations*, 14, p. 71-95.
- PIETRONI D., VAN KLEEF G. A., DE DREU C. K. W., PAGLIARO S. [2008], Emotions as strategic information : Effects of other's emotional expression on fixed-pie perception, demands, and integrative behavior in negotiation, *Journal of Experimental Social Psychology*, 44, p. 1444-1454.
- RABIN M. [1993], Incorporating Fairness into Game Theory and Econometrics, *American Economic Review*, 83, p. 1281-1302.
- REUBEN E., VAN WINDEN F. [2010], Fairness perceptions and prosocial emotions in the power to take, *Journal of Economic Psychology*, 31, p. 908-922.
- RICK S. et LOEWENSTEIN G. [2008], The Role of Emotion in Economic Behavior, in *Handbook of Emotions*, 3^e édition, LEWIS M., HAVILAND-JONES J., FELDMAN BARETT L. (eds), New York, Guilford, p. 138-156.
- ROBIN S. [2012], Mesure et caractérisation de l'attention à l'autre en situation d'interaction stratégique, *Revue de Philosophie Economique*, 13, p. 175-191.
- ROTEMBERG A. J. [2008], Minimally Acceptable Altruism and the Ultimatum Game, *Journal of Economic Behavior and Organization*, 66, p. 457-476.
- SALLY D. [2001], On sympathy and games, *Journal of Economic Behavior and Organization*, 44, p. 1-30.
- SANFEY A., RILLING J., ARONSON J., NYSTROM L., COHEN J. [2003], The neural basis of economic decision-making in the ultimatum game, *Science*, 300, p. 1755-1758.
- SCHNALL S., ROPER J., FESSLER D. M. T. [2010], Elevation leads to altruistic behavior, *Psychological Science*, 21, p. 315-320.
- SCHULTZ P. W., NOLAN J. M., CIALDINI R. B., GOLDSTEIN N. J., GRISKEVICIUS V. [2007], The Constructive, Destructive and Reconstructive Power of Social Norms, *Psychological Science*, 18, p. 429-434.
- SELIGMAN M. E. P., CSÍKSZENTMIHÁLYI M. [2000], Positive Psychology : An introduction, *American Psychologist*, 55, p. 5-14.
- SERRA D. [2007], Sentiments moraux et économie expérimentale, dans LIVET P., LEROUX A. (eds.), *Leçons de Philosophie Economique*, tome III (Sciences Economiques et Philosophie des Sciences), De Boeck Université.
- SINGER T., FEHR E. [2005], The Neuroeconomics of Mind Reading and Empathy, *American Economic Review Papers and Proceedings*, 95, p. 340-345.
- SMITH A. [1759], 1999, *Théorie des Sentiments Moraux*, Paris, Presses Universitaires de France.
- STAHL D. O., HARUVY E. [2006], Other-regarding preferences : Egalitarian warm glow, empathy, and group size, *Journal of Economic Behavior and Organization*, 61, p. 20-41.
- STARK O., FALK I. [1998], Transfers, Empathy Formation, and Reverse Transfers, *American Economic Review*, 88, p. 271-276.

- STRINGHAM E. P. [2011], Embracing morals in economics : The role of internal moral constraints in a market economy, *Journal of Economic Behavior and Organization*, 78, p. 98-109.
- SUGDEN R. [2002], Beyond sympathy and empathy : Adam Smith's concept of fellow-feeling, *Economics and Philosophy*, 18, p. 63-87.
- SUTTER M. [2007], Outcomes versus intentions : On the nature of fair behaviour and its development with age, *Journal of Economic Psychology*, 28, p. 69-78.
- TADELIS S. [2008], The power of shame and the rationality of trust, *Working paper series, center for responsible Business, UC Berkeley*.
- TANGNEY J. P., STUEWIG J., MASHEK D. J. [2007], Moral Emotions and Moral Behavior, *Annual review of psychology*, 58, p. 345-372.
- VÉLEZ GARCIA A. E., OSTROSKY-SOLIS F. [2006], From morality to moral emotions, *International journal of psychology*, 41, p. 348-354.
- WOLFESPERGER A. [2001], La modélisation économique de la rationalité axiologique. Des sentiments moraux aux mécanismes sociaux de la moralité, dans BOUDON R., DEMEULENAERE P., VIALE R. (dir.), *L'explication des normes sociales*, Paris, Presses Universitaires de France, p. 63-92.
- WOOD A. M., BROWN G. D. A., MALTBY J. [2011], Thanks, but I'm used to better : a relative rank model of gratitude, *Emotion*, 11, p. 175-180.
- WUBBEN M. J. J., DE CREMER D., VAN DIJK E. [2009], How emotion communication guides reciprocity : Establishing cooperation through disappointment and anger, *Journal of experimental social psychology*, 45, p. 987-990.
- XIAO E., HOUSER D. [2005], Emotion expression in human punishment behavior, *PNAS*, 102, p. 7398-7401.
- XIAO E., HOUSER D. [2009], Avoiding the sharp tongue : Anticipated written messages promote fair economic exchange, *Journal of Economic Psychology*, 30, p. 393-404.
- ZAK P. J. [2011a], The physiology of moral sentiments, *Journal of Economic Behavior and Organization*, 77, p. 53-65.
- ZAK P. J. [2011b], Moral Markets, *Journal of Economic Behavior and Organization*, 77, p. 212-233.