

**Philosophie expérimentale :
illustration par l'exemple de la punition morale**

Christine Clavien

(Département d'écologie et d'évolution, Université de Lausanne)

I. L'expérimentation se moque des barrières disciplinaires

Au XVII^e siècle, William Molyneux a formulé une question très pertinente pour la philosophie de l'esprit. Il s'agissait de savoir si un aveugle de naissance retrouvant soudainement la vue serait capable de distinguer, sans les toucher, une sphère d'un cube. Les plus grands philosophes de l'époque ont débattu de cette question au moyen de raisonnements logiques et d'expériences de pensées. Un siècle plus tard, les premières opérations de la cataracte ont fourni des éléments de réponse cruciaux. L'interprétation des données empiriques a encore longtemps fait débat¹ jusqu'à une récente étude scientifique dont le protocole expérimental suit scrupuleusement les conditions décrites par Molyneux². L'essentiel des résultats empiriques converge vers une réponse négative à la question originale : un minimum de pratique semble nécessaire pour le bon développement des capacités visuelles. Cet exemple montre qu'il peut être légitime et recommandable de recourir aux données expérimentales pour répondre à des interrogations philosophiques.

Un esprit sceptique pourrait suspecter que le problème de Molyneux est une exception à la règle. Au fond, il concerne nos facultés visuelles, dont les mécanismes sont infiniment plus simples que la plupart des objets de recherche philosophiques. Pensez notamment à toutes les thématiques relatives à la prise de décision humaine traitées en philosophie de l'action, philosophie morale ou philosophie politique. Les mécanismes et les influences causales impliqués sont si subtils et imbriqués qu'ils pourraient bien échapper à toute investigation expérimentale. Ainsi, on en viendrait à douter de la pertinence de l'approche empirique pour *la plupart* des débats

¹ Pour une revue, voir : M. Degenaar, *Molyneux's problem : three centuries of discussion on the perception of forms*, Dordrecht, Boston, Kluwer Academic Publishers, 1996.

² R. Held, Y. Ostrovsky, B. de Gelder, T. Gandhi, S. Ganesh, U. Mathur, P. Sinha, « The newly sighted fail to match seen with felt », in *Nature Neuroscience*, 14/5, 2011, pp. 551-53.

de nature philosophique. Mais un tel scepticisme est difficilement tenable. Reconnaissons pour commencer que les philosophes n'ont pas l'exclusivité des thématiques qu'ils traitent. Voici deux exemples illustratifs : les modalités de la prise de décision rationnelle sont beaucoup étudiées en économie³ et le rapport entre l'empathie et le comportement d'aide a été traité de manière systématique en psychologie⁴. Décréter que le comportement humain est trop compliqué pour être investigué empiriquement revient à rejeter la pertinence d'importants domaines académiques. Plus particulièrement, il faudrait affirmer que des résultats impressionnants tels que ceux de la fameuse expérience de Milgram⁵ sur la soumission à l'autorité ne nous disent rien d'intéressant sur le comportement humain⁶.

Une question plus délicate est de savoir s'il est judicieux que les philosophes eux-mêmes s'essayent à la pratique expérimentale dans l'objectif de faire avancer leurs raisonnements. Sachant qu'ils ne disposaient pas de connaissances et compétences médicales suffisantes, il est heureux que Molyneux, Locke ou Diderot ne se soient pas hasardés à opérer des patients souffrant de cataracte. Mais à supposer qu'ils aient fait preuve de telles connaissances et compétences, n'aurait-il pas été regrettable qu'ils se contraignent par peur de sortir de leur rôle de purs penseurs ?

Sur le principe, il semble donc que ce que l'on appelle aujourd'hui la « philosophie expérimentale » soit tout à fait recommandable à condition qu'elle soit pratiquée dans les règles de l'art. Précisons d'emblée que cette condition n'est pas aussi aisée à remplir qu'il n'y paraît et c'est probablement la raison pour laquelle elle suscite tant de scepticisme (souvent justifié) au sein de la communauté philosophique. En effet, la mise en œuvre d'une expérience scientifique est complexe, même lorsque les enjeux éthiques, le coût et les compétences techniques requises sont limités. Comme tout métier, l'expérimentation humaine est un art qui s'acquiert en théorie et en pratique. Il existe des filières académiques (notamment dans les départements de psychologie et d'économie) entièrement dédiées à cette

³ P. W. Glimcher, C. F. Camerer, E. Fehr, R. A. Poldrack, eds. *Neuroeconomics : decision making and the brain*. London, Academic Press, 2009.

⁴ C. D. Batson, *Altruism in humans*, Oxford, New York, Oxford University Press, 2011.

⁵ S. Milgram, E. Molinié, *Soumission à l'autorité : un point de vue expérimental*, Paris, Calmann-Lévy, 1974.

⁶ Une alternative pour couper toute relation entre l'expérimental et le philosophique serait de réduire drastiquement le champ d'investigation de ce dernier et le confiner à des questionnements auxquels il n'est pas pertinent (par ex. l'éthique normative) ou irréaliste (par ex. certains domaines métaphysiques) de chercher des réponses au moyen d'outils expérimentaux.

maîtrise. On y apprend comment poser des questions ou des hypothèses susceptibles d'être investiguées empiriquement, comment réfléchir à un modèle expérimental adéquat par rapport à l'objectif fixé⁷, comment analyser les résultats avec des outils statistiques appropriés, et comment interpréter ces résultats⁸.

Au-delà de cet obstacle de nature plutôt pratique il n'y a aucune raison de principe à fermer la porte de l'expérimentation aux philosophes. Au contraire, l'expérimentation s'avère être un puissant vecteur de discussion au sein et entre disciplines. Cette idée sera illustrée dans cet article au moyen d'un exemple. Il s'agira d'analyser l'impact théorique d'une étude sur la motivation à la punition morale. Cette étude⁹ a la particularité de se situer à la frontière entre la psychologie sociale et l'économie expérimentale tout en abordant des questions chères aux philosophes.

La section suivante pose le décor : elle expose les grandes lignes de (et met en relation) quatre débats interdisciplinaires impliquant des modèles explicatifs de la motivation morale. La section 3 ressert l'objectif en présentant un état des lieux des données expérimentales sur le volet punitif de la motivation morale. Les sections 4 et 5 sont consacrées à la description d'une étude particulière sur la punition morale et à l'analyse de son impact sur les modèles explicatifs généraux décrits dans la section 2.

II. Motivation morale – cadre théorique

Quels processus causaux sont impliqués dans l'action morale ? Cette question est de nature à intéresser les philosophes, mais également les

⁷ Entre autres, voici quelques éléments clefs qu'il importe de considérer lors de l'élaboration d'un modèle expérimental :

a) Si l'expérience est réalisée dans un milieu artificiel (par exemple en salle de cours avec des participants au courant de leur implication dans une expérience scientifique), il importe de réfléchir à la validité externe des résultats ; a-t-on de bonnes raisons de croire que les participants produiraient le même pattern de réponse dans leur environnement quotidien ?

b) Dans la mesure du possible, l'expérience doit être configurée de manière à éviter (ou du moins contrôler) l'effet de facteurs confondants (c'est-à-dire les facteurs causaux qui ne nous intéressent pas mais sont susceptibles d'influencer les résultats) ainsi que les corrélations et interactions entre plusieurs facteurs explicatifs.

c) Il est également crucial de s'assurer que la taille d'échantillon testé (en l'occurrence le nombre de participants à l'expérience) soit suffisante pour pouvoir appliquer les tests statistiques adéquats lors de l'analyse des données.

⁸ B. W. Pelham, H. Blanton, *Conducting research in psychology: measuring the weight of smoke*, Australia: Belmont, CA, Thomson Wadsworth, 2007.

⁹ C. Clavien, C. J. Tanner, F. Clément, M. Chapuisat, « Choosy moral punishers », in *Plos One*, 7/6, 2012, pp. e39002.

chercheurs issus d'autres disciplines telles que la psychologie, l'économie ou le droit. Elle est traitée dans au moins quatre débats historiques avec un certain nombre de nuances.

Le premier débat fait rage au sein de la philosophie et oppose les défenseurs de l'internalisme à ceux de l'externalisme. Selon la position internaliste, un jugement moral motive à l'action lorsqu'il est produit de manière sincère. Il existerait donc une relation causale nécessaire entre nos jugements et la motivation à produire les actions qu'ils prescrivent¹⁰. Par exemple, si nous pensons qu'il est moralement bon de partager un « don du ciel » (un bien inattendu que l'on reçoit sans contrepartie) avec les personnes de notre entourage qui n'ont pas été aussi chanceuses que nous, alors nous devrions, indépendamment de toute autre considération, ressentir le besoin de partager chaque fois que nous nous trouvons dans une situation de ce type. De même si nous pensons qu'il est moralement condamnable de profiter d'un bien commun après avoir refusé de contribuer à sa création, nous devrions être motivés à sanctionner les personnes qui agissent de la sorte. Selon la position externaliste en revanche, les jugements moraux ne possèdent pas de force motivante ; la relation entre le jugement moral et la motivation à l'action conforme à ce jugement est de nature externe et contingente. Ainsi, nous aidons nos voisins et sanctionnons les profiteurs pour des raisons non morales (pour éviter de soulever des réactions de désapprobation de la part de nos partenaires sociaux, par peur de l'œil inquisiteur de Dieu, etc.).

Un débat connexe, à l'interface entre la philosophie et la psychologie, oppose les défenseurs de certaines versions de l'éthique de la vertu¹¹ aux situationnistes¹². Le premier groupe de penseurs soutient que l'action morale découle d'un caractère vertueux et que les traits de caractères vertueux (générosité, honnêteté, etc.) sont des dispositions profondes, constantes et stables qui peuvent être éduquées et développées¹³.

¹⁰ D. O. Brink, *Moral realism and the foundations of ethics*, Cambridge, New York, Cambridge University Press, 1989.

¹¹ R. Hursthouse, *On virtue ethics*, Oxford ; New York, Oxford University Press, 1999.

¹² G. Harman, « Moral philosophy meets social psychology: virtue ethics and the fundamental attribution error », in *Proceedings of the Aristotelian Society*, 99, 1999, pp. 315-31 ; J. M. Doris, *Lack of character*, Cambridge, Cambridge University Press, 2002.

¹³ Il existe diverses formes d'éthiques de la vertu. Certaines n'intègrent pas explicitement la notion de trait de caractère. D'autres l'intègrent mais sous forme « idéaliste » : les traits de caractère vertueux sont des idéaux normatifs quasiment inatteignables par le commun des mortels, mais ils servent d'exemples, guident nos actions et nous aident à nous développer. De par leur caractère normatif et idéaliste, ces positions ne peuvent pas être soumises à des observations empiriques qui impliquent des gens ordinaires. Pour cette raison, cet article

Par opposition, les situationnistes doutent de la pertinence de ce modèle causal. Pour eux, plutôt que les traits de caractère, ce sont les circonstances propres aux situations dans lesquelles se trouvent les personnes au moment de la prise de décision (événement stressant ou agréable, pression sociale, alternatives d'actions possibles, etc.) qui prédisent comment elles vont agir. Précisons que le situationnisme n'est pas incompatible avec l'idée que certains traits de caractères influencent nos actions¹⁴. En revanche, ses défenseurs doutent que les traits spécifiques généralement attribués aux personnes vertueuses (courage, honnêteté, bienveillance, amabilité) ou vicieuses exercent un pouvoir causal digne d'être pris en considération.

Quoique ces deux débats ne soient pas souvent mis en relation directe, un défenseur de l'éthique de la vertu est généralement partisan de l'internalisme. Il s'agit d'une question de cohérence : une personne motivée par un trait de caractère vertueux est en principe convaincue que ses actions sont moralement requises. L'éthique de la vertu peut donc être considérée comme une version précise et concrète de l'internalisme : elle propose une hypothèse spécifique sur le mécanisme (le caractère vertueux) responsable de la relation causale entre le jugement et l'action morale¹⁵. Par opposition, le situationnisme, qui fait appel aux effets du contexte pour expliquer l'action morale, conduit à l'externalisme.

Le troisième débat a majoritairement impliqué les philosophes et les psychologues. Il s'agit de savoir si les êtres humains sont capables d'actions altruistes ou s'ils agissent toujours de manière égoïste. Dans ce contexte, une action est dite altruiste si elle découle directement et uniquement d'une motivation (désir, réaction émotionnelle, impulsion) à améliorer ou contribuer au bien-être d'autrui. Les défenseurs de l'altruisme

traite uniquement des versions non idéalistes d'éthique de la vertu, et parmi elles, de celles qui intègrent la notion de caractère.

¹⁴ Même un situationniste aussi radical que Harman (1999) admet le pouvoir causal de traits généraux comme la timidité ou la jovialité (quoiqu'il préfère les qualifier d'« aspects innés du tempérament » plutôt que de « caractères », p. 316).

¹⁵ On pourrait imaginer qu'une personne agisse sous l'impulsion d'un caractère vertueux sans pour autant poser un jugement moral conforme à cette action. Dans ce cas de figure, le caractère vertueux court-circuiterait en quelque sorte le jugement moral. Mais cette nuance logique est de moindre conséquence pour notre propos. D'une part, cette possibilité n'empêche pas que dans les situations où cette même personne vertueuse prend le temps de poser un jugement moral, son caractère vertueux l'incite à agir dans le même sens. D'autre part, les cas d'inconsistance entre les jugements moraux et l'expression du caractère vertueux sont généralement ignorés ou traités comme des exceptions par les défenseurs de l'éthique de la vertu.

psychologique¹⁶ répondent par la positive alors que les défenseurs de l'égoïsme psychologique¹⁷ font preuve de scepticisme à cet égard. Ces derniers pensent que toute action apparemment altruiste est en réalité causée par des motifs dirigés vers soi-même (le désir de se forger une bonne réputation, la peur de la punition divine, etc.).

Le rapport entre l'altruisme et la morale ne peut pas être établi sans une prise de position précise sur la nature de la moralité¹⁸ même s'il est vrai que beaucoup de penseurs conçoivent le désir d'améliorer ou de contribuer au bien-être d'autrui comme une marque de bonne moralité. A première vue, on pourrait penser que l'altruiste psychologique mène à l'internalisme et à l'éthique de la vertu. Ce rapport n'est cependant pas évident car la thèse de l'altruisme psychologique est ciblée. Elle exige uniquement la possibilité qu'une motivation dirigée vers autrui cause une action généreuse. En soi, cela n'implique pas que cette motivation révèle une relation constante entre jugement et action dans tous les autres domaines de la morale (condition nécessaire à l'internalisme) ou que cette motivation relève d'une disposition vertueuse profonde et stable (condition nécessaire à l'éthique de la vertu). De même, les implications de la position égoïste psychologique ne sont pas toutes évidentes : l'acceptation de l'externalisme et le rejet de l'éthique de la vertu semblent clairement en découler mais les liens avec la position situationniste restent troubles puisque cette dernière s'oppose également à l'idée de caractères non vertueux.

Un quatrième débat, qui évolue à l'interface entre l'économie et d'autres sciences sociales telles que la sociologie ou la psychologie, partage les origines historiques de la controverse entre l'égoïsme et l'altruisme. Il s'agit d'évaluer la pertinence du modèle (très caricatural) de l'*homo economicus* selon lequel les êtres humains agissent toujours de manière à maximiser leur utilité, cette dernière se réduisant à des biens égoïstes tels que le plaisir individuel ou l'argent¹⁹. Dans cette littérature, beaucoup

¹⁶ J. Butler, « Fifteen sermons », in *British moralists, 1650-1800 : selected and edited with comparative notes and analytical index*, Oxford, Clarendon Press, 1991, pp. 325-77 ; C. D. Batson, « Moral masquerades: Experimental exploration of the nature of moral motivation », in *Phenomenology and the Cognitive Sciences*, 7/1, 2008, pp. 51-66.

¹⁷ R. B. Cialdini, M. Schaller, D. Houlihan, K. Arps, J. Fultz, A. L. Beaman, « Empathy-based helping: Is it selflessly or selfishly motivated? », in *Journal of Personality and Social Psychology*, 52/4, 1987, pp. 749-58 ; B. Mandeville, *La fable des abeilles : première partie*, Paris, Vrin, 1998 (1729).

¹⁸ C. D. Batson, *Altruism in humans*, Oxford, New York, Oxford University Press, 2011 ; C. Clavien, « Trois sortes d'altruisme et leur rapport à la morale », in *La morale humaine et les sciences*, Paris, Editions Matériologiques, 2011, pp. 141-68.

¹⁹ H. A. Simon, *The sciences of the artificial*, Cambridge, Mass., MIT Press, 1996 (1969) ; P. Bourdieu, *Les structures sociales de l'économie*, Paris, Seuil, 2000 ; E. Fehr, U.

d'auteurs sont davantage intéressés à développer une formule efficace de la fonction d'utilité (notion fondamentale en économie théorique) plutôt qu'à définir de manière précise les mécanismes causaux sous-jacents aux actions généreuses. Malgré cela, une théorie causale assez particulière a été développée au sein de l'économie expérimentale²⁰. Il s'agit de l'hypothèse de la « réciprocité forte » (ci-après, RF), selon laquelle les êtres humains possèdent, d'une part une tendance à coopérer et se conformer aux normes en vigueur dans leur société (volet positif de la RF), et d'autre part une répulsion face aux tricheurs, profiteurs et transgresseurs des normes sociales, doublée d'un désir de les punir (volet négatif de la RF).²¹ Selon cette hypothèse, la RF est un trait de caractère réparti de manière universelle au sein de l'espèce humaine mais qui s'exprime de manière plus ou moins prégnante selon les types d'individus (de même que l'on trouve des personnes plus ou moins timides au sein de chaque population). Pour des raisons qui gagneraient à être décryptées par les historiennes ou sociologues des sciences, cette hypothèse a été largement discutée au-delà du domaine économique, notamment en anthropologie évolutionnaire²², en philosophie²³, en psychologie sociale et en droit²⁴. Certains auteurs qui s'inspirent des recherches en économie expérimentale sont d'ailleurs extrêmement explicites sur les mécanismes psychologiques impliqués dans

Fischbacher, « The nature of human altruism », in *Nature*, 425/6960, 2003, pp. 785-91 ; G. Gigerenzer, *Gut feelings: short cuts to better decision making*, London, New York, Penguin Books, 2008 (2007) ; E. Fehr, C. Camerer, « Social neuroeconomics: the neural circuitry of social preferences », in *Trends in Cognitive Sciences*, 11/10, 2007, pp. 419-27 ; H. Gintis, J. Henrich, S. Bowles, R. Boyd, E. Fehr, « Strong reciprocity and the roots of human morality », *Social Justice Research*, 21/2, 2008, pp. 241-53.

²⁰ J. P. Carpenter, P. H. Matthews, O. Ong'ong'a, « Why punish? Social reciprocity and the enforcement of prosocial norms », in *Journal of Evolutionary Economics*, 14/4, 2004, pp. 407-29 ; E. Fehr, U. Fischbacher, « Third-party punishment and social norms », in *Evolution and Human Behavior*, 25/2, 2004, pp. 63-87 ; B. Rockenbach, M. Milinski, « The efficient interaction of indirect reciprocity and costly punishment », in *Nature*, 444/7120, 2006, pp. 718-23 ;

²¹ Notons que le volet négatif de la RF est généralement associé à l'idée de coût que la personne est prête à payer pour pratiquer la punition morale.

²² E. Fehr, J. Henrich, « Is strong reciprocity a maladaptation? On the evolutionary foundations of human altruism », in *Genetic and cultural evolution of cooperation*, Cambridge, Mass., MIT Press, 2003, pp. 55-82.

²³ Voir l'analyse critique de : F. Guala, « Reciprocity: weak or strong? What punishment experiments do (and do not) demonstrate », in *Behavioral and Brain Sciences*, 35/1, 2012, pp. 1-15.

²⁴ K. M. Carlsmith, J. M. Darley, P. H. Robinson, « Why do we punish? Deterrence and just deserts as motives for punishment », in *Journal of Personality & Social Psychology*, 83/2, 2002, pp. 284-99.

le volet punitif de la RF. On peut lire chez Darley²⁵ ou Carlsmith et collègues²⁶ que lorsqu'une personne prend conscience d'une transgression envers elle-même ou envers autrui, elle fait l'expérience d'une réaction intuitive et émotionnelle d'outrage moral, elle blâme cette action et cherche à appliquer une punition ajustée au degré de ce blâme²⁷. Dans la littérature, cette hypothèse explicative est désignée par le terme de « just desert theory » (que l'on pourrait traduire par théorie de la punition méritée).

A défaut d'être discutée dans le détail, la parenté entre la RF et la moralité est explicitement revendiquée par ses défenseurs²⁸. En outre, même s'ils ne s'expriment pas directement sur cette question, ces auteurs (ainsi que les défenseurs de la « just desert theory ») semblent se ranger dans le camp de l'internalisme moral et de l'éthique de la vertu. Plus précisément, ils se prononcent explicitement sur un trait de caractère moral (la réciprocité forte) qui guide les personnes vertueuses dans leurs actions.

Quels que soient les termes utilisés et les nuances de sens propres à chacun de ces débats, les penseurs et chercheurs impliqués se demandent (ou contribuent à la question de savoir) si nos actions morales sont intrinsèquement liées à nos convictions (jugements, croyances) morales ou si des éléments extrinsèques (calcul d'intérêt personnel, pression externe, état d'esprit de nature amoral dans lequel on se trouverait) sont nécessaires pour nous motiver à agir.

III. Punition morale – données expérimentales

²⁵ J. M. Darley, « Morality in the law: the psychological foundations of citizens' desires to punish transgressions », in *Annual Review of Law and Social Science*, 5/1, 2009, pp. 1-23.

²⁶ K. M. Carlsmith, J. M. Darley, P. H. Robinson, « Why do we punish? Deterrence and just deserts as motives for punishment », in *Journal of Personality & Social Psychology*, 83/2, 2002, pp. 284-99.

²⁷ « When a person registers a transgression against self or others, the person experiences an intuitively produced, emotionally tinged reaction of moral outrage. The reaction is driven by the just deserts-based retributive reactions of the person. » (Darley, 2009, p. 1)

« Intentional harms call forth sanctions. (...) The observer experiences some degree of moral outrage, is able to assign a degree of blame for that act, and then generates a punishment commensurate with the blame. (...) Unless the punishment is imposed, a real feeling of incompleteness lingers, and there is a sense that justice has not been done. These feelings of incompleteness and sense of failed justice are held by those who witness or become aware of the original offense as well as by its victim. » (Carlsmith, et al., 2002, pp. 296-97)

²⁸ H. Gintis, S. Bowles, R. Boyd, E. Fehr, éd. *Moral sentiments and material interests : the foundations of cooperation in economic life*. Cambridge, Mass., MIT Press, 2005 ; H. Gintis, J. Henrich, S. Bowles, R. Boyd, E. Fehr, « Strong reciprocity and the roots of human morality », in *Social Justice Research*, 21/2, 2008, pp. 241-53.

Dans le domaine expérimental, le volet positif de la motivation morale, c'est-à-dire l'aide et la bienveillance, compte de très belles études²⁹. En revanche, le versant négatif, c'est-à-dire la punition morale, demeure peu exploré³⁰. Il existe bien des études sur les facteurs qui influencent les décisions judiciaires en situation réelle³¹, mais les résultats sont peu révélateurs du comportement punitif des gens au quotidien puisque les juges ont pour tâche (rémunérée) d'appliquer la punition selon des critères établis. Plus pertinente pour notre propos, il existe depuis quelques années une ligne de recherche en économie expérimentale sur la « punition altruiste », laquelle correspond au versant négatif de la RF³². Ces études impliquent généralement des étudiants d'université qui reçoivent une somme d'argent à investir (ou ne pas investir) au cours de jeux pratiqués sur ordinateurs. Les règles des jeux donnent la possibilité aux participants d'effectuer des choix

²⁹ Notamment : B. Latane, J. M. Darley, *The unresponsive bystander - Why doesn't he help?*, Appleton-Century Crofts, 1970, 1970 ; A. M. Isen, P. F. Levin, « Effect of feeling good on helping - cookies and kindness », in *Journal of Personality and Social Psychology*, 21/3, 1972, pp. 384-388 ; J. M. Darley, C. D. Batson, « "From Jerusalem to Jericho": A study of situational and dispositional variables in helping behavior », in *Journal of Personality and Social Psychology*, 27/1, 1973, pp. 100-08 ; R. B. Cialdini, M. Schaller, D. Houlihan, K. Arps, J. Fultz, A. L. Beaman, « Empathy-based helping: Is it selflessly or selfishly motivated? », in *Journal of Personality and Social Psychology*, 52/4, 1987, pp. 749-58 ; R. A. Baron, « A whiff of reality: Positive affect as a potential mediator of the effects of pleasant fragrances on task performance and helping », in *Environment and Behavior*, 26/6, 1994, pp. 766-84 ; C. D. Batson, « Moral masquerades: Experimental exploration of the nature of moral motivation », in *Phenomenology and the Cognitive Sciences*, 7/1, 2008, pp. 51-66.

³⁰ Il existe une importante littérature sur la condamnation morale (voir notamment : J. D. Greene, R. B. Sommerville, L. E. Nystrom, J. M. Darley, J. D. Cohen, « An fMRI investigation of emotional engagement in moral judgment », in *Science*, 293/5537, 2001, pp. 2105-08 ; J. Haidt, « The emotional dog and its rational tail: A social intuitionist approach to moral judgment », in *Psychological Review*, 108/4, 2001, pp. 814-34 ; S. Nichols, « After objectivity: an empirical study of moral judgment », in *Philosophical Psychology*, 17, 2004, pp. 3-26 ; C. R. Sunstein, « Moral heuristics », in *Behavioral and Brain Sciences*, 28, 2005, pp. 531-42 ; R. L. Woolfolk, J. M. Doris, J. M. Darley, « Identification, situational constraint, and social cognition: Studies in the attribution of moral responsibility », in *Cognition*, 100/2, 2006, pp. 283-301 ; J. Haushofer, E. Fehr, « You shouldn't have: Your brain on others' crimes », in *Neuron*, 60/5, 2008, pp. 738-40 ; J. M. Knobe, S. Nichols, *Experimental philosophy*, Oxford, New York, Oxford University Press, 2008 ; S. Danziger, J. Levav, L. Avnaim-Pesso, « Extraneous factors in judicial decisions », in *Proceedings of the National Academy of Sciences*, 108/17, 2011, pp. 6889-92), mais il est difficile d'exploiter ces résultats pour la question qui nous occupe puisqu'il n'est pas certain qu'un jugement soit un facteur motivant à l'action.

³¹ I. V. Blair, C. M. Judd, K. M. Chappleau, « The influence of Afrocentric facial features in criminal sentencing », in *Psychological Science*, 15/10, 2004, pp. 674-79.

³² E. Fehr, U. Fischbacher, « Third-party punishment and social norms », in *Evolution and Human Behavior*, 25/2, 2004, pp. 63-87 ; J. P. Carpenter, P. H. Matthews, « Norm enforcement: anger, indignation, or reciprocity? », in *Journal of Economic Behavior & Organization*, 10/3, 2012, pp. 555-72.

injustes (par exemple, partager une somme d'argent de manière inéquitable avec un autre participant) et, dans une étape ultérieure du jeu, de punir les participants qui ont opté pour ces choix injustes.

Les expériences montrent que les gens sont enclins à appliquer la punition lorsqu'ils ont eux-mêmes été victimes d'une injustice³³ ou lorsqu'ils peuvent se forger une réputation de justicier aux yeux de leurs congénères³⁴. Mais ces situations, partiellement interprétables en termes d'intérêts personnels (vengeance, réputation), sont peu révélatrices pour notre questionnement. Ce qui nous intéresse davantage au regard de la motivation morale et des débats autour de l'internalisme et du situationnisme, est de savoir si les gens sont motivés à punir les infractions morales dans des situations où ils n'ont aucune motivation externe (notamment rien à y gagner personnellement) à pratiquer cette punition. Il importe donc d'explorer les situations dans lesquelles le seul motif plausible pour la punition est le désir de rétablir la justice ou de renforcer une norme morale. Si une tendance à punir dans ce type de situations est révélée par les études empiriques, les positions de type internaliste gagnent en crédibilité.

Une sous-classe de jeux pratiqués en économie expérimentale a été spécialement pensée pour aborder cette problématique. Au cours de ces jeux, une partie des participants sont placés en situation d'observateurs (« third party ») anonymes ; ils ne sont pas eux-mêmes impliqués dans le jeu d'argent mais spectateurs des interactions entre d'autres joueurs ; ils ont la possibilité (mais aucune obligation) de punir, moyennant investissement d'une certaine somme ; ni les autres joueurs affectés par leur décision, ni les expérimentateurs ne connaissent leur décision individuelle (pas de traçabilité de l'origine de la punition). Les résultats obtenus dans ces expériences montrent une légère tendance à appliquer la punition³⁵.

Il semblerait donc que nous disposions de bonnes raisons empiriques de soutenir l'internalisme (ou l'éthique de la vertu, la RF ou la « just desert theory ») face à ses détracteurs que sont l'externalisme et le situationnisme. Quelques doutes subsistent toutefois. Un certain nombre de données issues

³³ D. J. F. de Quervain, U. Fischbacher, V. Treyer, M. Schellhammer, U. Schnyder, A. Buck, E. Fehr, « The neural basis of altruistic punishment », in *Science*, 305/5688, 2004, pp. 1254-58.

³⁴ P. Barclay, « Reputational benefits for altruistic punishment », in *Evolution and Human Behavior*, 27/5, 2006, pp. 325-44.

³⁵ R. Kurzban, P. DeScioli, E. O'Brien, « Audience effects on moralistic punishment », in *Evolution and Human Behavior*, 28/2, 2007, pp. 75-84 ; P. G. Lewisch, S. Ottone, F. Ponzano, « Free-riding on altruistic punishment? Experimental comparison of third-party-punishment in a stand-alone and in an in-group environment », in *Review of Law & Economics*, 7/1, 2011, pp. 165-94.

de l'anthropologie et de la sociologie indiquent que la punition est uniquement appliquée lorsque des pressions sociales ou de l'intérêt personnels sont en jeu³⁶. De plus, les études menées par les économistes expérimentaux sur le comportement social humain s'effectuent généralement dans un contexte artificiel (en laboratoire, avec des participants conscients d'être engagés dans une expérience scientifique) à forte connotation économique. On peut se demander si ce genre de contexte est approprié pour tester la motivation morale. De manière générale, ces études sont confrontées à la difficulté de la validité externe des résultats obtenus ; peut-on sérieusement penser que les individus qui ont appliqué la punition morale dans de telles conditions se comporteront de même dans leur quotidien ? La meilleure manière de s'en rendre compte est d'imaginer un modèle expérimental similaire (participants placés en observateurs anonymes qui ont la possibilité de punir une infraction morale) mais qui surprend les participants en situation réelle, dans leur environnement quotidien. Cela a été réalisé dans une étude récente sur la punition morale³⁷ qui sera décrite et discutée en détail dans les deux sections suivantes.

IV. Etude sur la punition morale – modèle expérimental

L'expérience a lieu dans une salle informatique d'un établissement scolaire, durant un cours ordinaire. Les participants (étudiants de l'établissement scolaire) ne se doutent pas qu'ils sont impliqués dans une étude scientifique. Les expérimentateurs se présentent comme des représentants d'une nouvelle maison de disque en phase de lancement, qui organise un concours de violon auxquels participent des violonistes en fin de formation. La maison de disque s'engage à enregistrer un morceau solo joué par le/la gagnant/e du concours et le diffuser largement. Il s'agit donc d'une importante promotion pour un/une jeune violoniste en début de carrière. Les représentants de la maison de disque demandent aux étudiants-participants de prendre le rôle de juges pour la finale de ce concours. Leur tâche est de visionner sur ordinateur la prestation filmée de deux filles présélectionnées par un premier jury et de voter pour une de ses deux

³⁶ Pour une revue de cette littérature, voir : F. Guala, « Reciprocity: weak or strong? What punishment experiments do (and do not) demonstrate », in *Behavioral and Brain Sciences*, 35/1, 2012, pp. 1-15.

³⁷ C. Clavien, C. J. Tanner, F. Clément, M. Chapuisat, « Choosy moral punishers », in *Plos One*, 7/6, 2012, pp. e39002.

violonistes.³⁸ En plus de la prestation musicale, les participants visionnent l'interview d'un ancien professeur de ces violonistes. Le professeur parle des qualités techniques et sociales des deux musiciennes. Après visionnement des films, les étudiants votent de manière anonyme et répondent à quelques questions d'ordre général. Cela fait, ils enregistrent leurs réponses sur une clef USB non labellisée, la déposent dans un grand sac commun et sortent de la salle.

Les séquences de musiques ont été choisies de manière à ce que la prestation musicale d'une des deux concurrentes surpasse celle de la seconde (selon des critères professionnels). De plus, la moitié des participants-juges, reçoivent des informations sociales positives sur les deux violonistes par le biais du visionnement de l'interview (condition neutre). Le professeur dit des deux filles qu'elles sont sympathiques, bien intégrées dans la classe et appréciées de tout le monde. En revanche, l'autre moitié des participants-juges apprennent qu'une des deux musiciennes (celle qui joue objectivement mieux) a un caractère moralement répugnant (condition test). Le professeur dit qu'elle est mal insérée dans l'école pour cause de troubles de comportement récurrents. A titre d'illustration, il mentionne deux exemples de ses méfaits : elle a mélangé la partition d'une collègue juste avant un concert, ce qui a eu pour effet de déstabiliser la « victime de cette machination » et mis tout l'orchestre dans un état de stress ; une autre fois, elle a désaccordé un instrument avant un concert ce qui a eu le même effet déstabilisant. En bref, dans cet expérience, une information morale est rendue saillante pour la moitié des participants et non saillante pour l'autre moitié des participants.

Ce dispositif permet de tester la propension des gens à appliquer la punition morale. Plus précisément, il permet d'observer dans quelle mesure un jugement moral peut influencer, voire contrebalancer l'effet causal d'un jugement esthétique : si tel est le cas, les participants-juges de la condition test se montrent moins enclin à voter pour la violoniste talentueuse mais immorale. Ce choix manifeste d'une forme de punition morale³⁹ puisqu'il y

³⁸ Pour éviter tout biais de choix, les violonistes sont physiquement similaires, habillées en noir et leurs visages sont floutés dans la projection filmée.

³⁹ Deux éléments sont à relever par rapport à cette notion de punition morale. Premièrement, on peut discuter de son caractère « coûteux ». Il est certain que les participants-juges ne doivent pas investir une somme d'argent (comme c'est le cas dans les jeux en économie expérimentale) pour punir la violoniste immorale. En revanche, ils doivent renoncer à leur évaluation esthétique ce qui implique un certain coût psychologique. Deuxièmement, il n'est pas certain que tous les participants influencés par l'information morale conçoivent consciemment leur décision de voter pour la violoniste moins talentueuse comme étant une punition morale ; pour certains le caractère moral

a refus de contribuer à la carrière de cette fille au profit d'une violoniste moins talentueuse mais moralement correcte⁴⁰.

Le fait que les participants-juges soient observés dans leur contexte scolaire habituel, sans savoir qu'ils participent à une expérience scientifique, crédite les résultats observés d'une forte validité externe ; les effets parasites liés à un contexte expérimental artificiel sont négligeables. De plus, les participants-juges se trouvent en position d'observateurs anonymes et n'ont aucune perspective d'interaction future avec les personnes affectées par leur décision d'action. Ce dispositif permet d'éviter des facteurs confondants qui pervertissent l'interprétation de nombreuses autres études. En effet, toutes les pressions sociales liées à la réputation ou à l'interaction avec les personnes impliquées dans la situation conflictuelle peuvent être considérés comme négligeables.

V. Etude sur la punition morale – résultats et interprétation

Le test a été mené sur trois catégories socio-professionnelles. Des étudiants de gymnase (équivalent du bac français), des étudiants de Haute Ecole Pédagogique (école préparatoire au métier d'enseignant), des étudiants de l'académie de police (aspirants policiers). Les résultats sont surprenants. Seuls les futurs enseignants ont puni la jeune violoniste talentueuse mais immorale : ils se sont montrés moins enclins à voter pour elle (de manière statistiquement significative) lorsqu'ils étaient informés de son mauvais comportement envers ses collègues violonistes. En revanche, l'information morale n'a eu strictement aucun impact sur le vote des gymnasiens et des aspirants policiers ; ces derniers ne semblent pas avoir ressenti ou considéré l'aspect moral comme pertinent pour leur décision de vote.

Au vu de ces résultats, deux questions vont nous occuper dans cette section. Quel est l'impact de cette recherche sur les débats autour de la motivation morale? Comment expliquer que seuls les enseignants se soient montrés sensibles à l'information morale ?

pourrait être conçu comme un ingrédient nécessaire à l'évaluation esthétique. Il importe ici de concevoir la punition morale dans un sens faible, l'idée étant que l'ingrédient moral entre en ligne de compte dans la prise de décision des participants-juges.

⁴⁰ Au cours de l'expérience, les participants croient réellement que leur décision de vote aura un impact sur l'avenir professionnel des violonistes ; cela permet de considérer leur vote comme une action plutôt qu'un simple jugement.

L'internalisme et la théorie de la vertu évoquent le plus souvent la motivation à aider autrui. Mais en principe, ces théories ont des implications plus larges. Si l'on est vertueux ou motivé à agir en conformité avec nos jugements moraux, on devrait non seulement être incité à aider les vieilles dames à traverser la route, mais également à sanctionner ceux qui empêchent ostensiblement les vieilles dames de traverser la route. Plus précisément, on s'attend à ce qu'une personne réellement vertueuse, ou motivée par ses jugements moraux, applique la punition morale, et cela même en l'absence de pression externe ou d'intérêt personnel. Cette idée est défendue de manière explicite par les défenseurs de la RF, et de la « just desert theory ». Ainsi, si aucune clause spécifique n'y est ajoutée, toutes ces théories prédisent que l'on devrait observer au sein de n'importe quelle population humaine, une certaine propension (qu'on la nomme « trait de caractère », « vertu », « disposition » ou « tendance comportementale ») à punir les comportements qui sont généralement considérés comme immoraux au sein de la population en question. Il est clair que cette propension peut varier selon le nombre de personnes vertueuses présentes, selon leur degré de leur vertu et selon la nature de l'infraction morale ; mais si cette propension est supposée être une propriété de la nature humaine, elle doit pouvoir être observée, via certaines constances de comportement, dans tout groupe de personnes suffisamment large.

Or, dans l'étude sur la punition décrite ci-dessus, une transgression morale perpétrée par une violoniste (nuisances répétées et gratuites contre des collègues étudiants) a fait réagir une catégorie de participants (les enseignants) mais pas les deux autres (les gymnasiens et les policiers). Le problème ne vient pas tellement du fait que deux catégories n'aient pas réagi. En effet, les théories de type internaliste prédisent seulement que les gens ont tendance à punir, mais cette tendance peut être réprimée par d'autres facteurs ; par exemple la force motivante du jugement esthétique ou un souci d'objectivité (on peut penser qu'un juge dans un concours de musique se doit de faire abstraction des aspects non musicaux). Le problème vient davantage de la réponse différenciée entre les catégories socio-professionnelles testées. En effet, tous les participants à l'expérience proviennent du même pays et de la même région linguistique. Globalement, ils ont donc une base socio-culturelle commune. De plus, ils sont tous placés dans une situation identique (dans la salle informatique de leur école, recevant les mêmes informations, devant réaliser la même tâche) et ne subissent aucune pression de la part des expérimentateurs ou de leurs collègues (puisque le vote est strictement anonyme). Ainsi, si l'on trouve de

la punition morale chez les enseignants, l'internalisme (tout comme la théorie de la vertu, la RF ou la « Just desert theory ») prédit que l'on devrait trouver le même pattern de comportement chez les gymnasiens et les policiers, car il n'y a aucune raison de penser que les membres de ces deux catégories n'aient pas perçu le caractère immoral de la violoniste talentueuse⁴¹. Or l'expérience révèle le contraire.

Les théories de type internaliste sont clairement ébranlées par ces résultats expérimentaux. Faut-il les abandonner au profit d'une position externaliste ou situationniste ? Pas forcément car face à des résultats empiriques contradictoires avec une théorie, trois voies sont possibles pour maintenir cette dernière.

La première voie, très saine si elle est utilisée sans *à priori* négatif contre l'expérimentation en général, consiste à questionner la valeur scientifique de l'expérience. Il se peut que des tests statistiques non appropriés aient été utilisés, ou qu'un aspect crucial de la procédure expérimentale ait été mal maîtrisé. Dans certains cas il est même permis de suspecter les expérimentateurs d'avoir inventé ou trafiqué leurs résultats. Face à chaque expérience scientifique, il est important de se poser ces questions. Mais dans le même temps, mettre en cause la valeur scientifique d'une étude est une accusation grave ; c'est pourquoi elle doit être étayée. Si aucune erreur flagrante ne peut être imputée aux expérimentateurs, on peut demander (ou tenter soi-même) de répliquer les données obtenues. C'est d'ailleurs une procédure extrêmement courante et importante en science. Un résultat expérimental acquiert une réelle crédibilité au sein de la communauté scientifique lorsqu'il a été reproduit à plusieurs reprises, idéalement par des laboratoires indépendants.

Si l'on fait confiance à la qualité scientifique de l'étude, une seconde voie se dessine. Il s'agit de proposer des explications alternatives pour les résultats obtenus, qui soient compatibles avec la théorie mise en question. Pour le cas qui nous occupe, il serait très peu crédible de recourir à l'échappatoire de la variété des valeurs morales défendues dans des groupes sociaux distincts ; les policiers et les gymnasiens admettent certainement qu'il est immoral (ou du moins socialement blâmable) de saboter le travail des autres sans raison. En revanche, on pourrait émettre l'hypothèse que les enseignants sont plus vertueux (influencés dans leur comportement par les

⁴¹ Une telle hypothèse est très peu crédible au vu des commentaires laissés par les participants dans les formulaires de vote. Plus de 40% des policiers et gymnasiens font explicitement référence au caractère immoral de la violoniste ; la plupart expliquent qu'ils ne veulent pas se baser sur cette information pour leur choix de vote.

informations morales) que les autres catégories socio-professionnelles, ou que la vertu morale ne s'acquière qu'avec une certaine maturité intellectuelle ou une capacité d'attention dont ne font pas preuve les gymnasiens et les policiers⁴². Ces derniers pourraient bien percevoir l'immoralité de la violoniste mais se montreraient incapables de la juger 'sincèrement' comme immorale, d'où leur manque de motivation à punir⁴³. Une hypothèse alternative serait de dire que les enseignants se montrent moins aptes à réprimer leur jugement moral que les deux autres catégories socio-professionnelles, lesquelles sont plus objectives dans leur décision de vote (sous-entendu qu'un juge de musique ne doit pas se laisser distraire par d'autres critères)⁴⁴. Remarquons toutefois que ces interprétations des résultats sont lourdes de conséquences : dans le premier cas, on impute un manque de réceptivité aux données morales, de maturité intellectuelle ou de capacité d'attention aux gymnasiens et aux policiers ; dans le deuxième cas, on impute un déficit d'objectivité aux enseignants. En outre, ces hypothèses requièrent des explications supplémentaires : on se demande quels facteurs sont responsables d'une telle répartition des compétences (moralité, maturité intellectuelle, faculté d'attention, objectivité) entre les différentes catégories socio-professionnelles. A cet égard, deux options se dessinent : la qualité ou la nature de l'éducation fournie par les différentes écoles, ou une tendance des personnes dotées de certaines compétences à choisir un certain type de carrière professionnelle⁴⁵.

Si l'on accepte la validité scientifique de l'étude ainsi que son interprétation problématique pour la théorie mise en question, une troisième voie se dessine. Elle consiste à proposer des révisions à la théorie de sorte à la rendre compatible avec les données expérimentales. Dans cette optique, une solution pour les approches de type internaliste, serait de supposer que la motivation à la punition morale s'applique en relation avec certaines

⁴² Précisons que l'âge n'est pas un critère plausible puisque les catégories "futurs enseignants" et "futurs policiers" sont composées d'individus d'âge comparable.

⁴³ Une autre manière de formuler cette idée serait de dire que les enseignants exigent des violonistes qu'elles fassent preuve de bonne conduite alors que les policiers et gymnasiens n'ont pas de telle attente par rapport aux violonistes. De ce fait, ces derniers restent froids face à l'acte immoral ; ils ne ressentent pas le sentiment d'outrage décisif pour motiver à punir.

⁴⁴ Une alternative consisterait à proposer un mélange de ces deux hypothèses. On pourrait supposer que les gymnasiens manquent de maturité intellectuelle et que les policiers font preuve d'une objectivité particulière puisqu'ils ont été éduqués à réaliser les tâches qui leurs sont données sans interférence avec leurs jugements personnels.

⁴⁵ Notons que le gymnase est une école préparatoire à toutes les hautes écoles (y compris la formation de futurs enseignants). On ne peut donc pas attribuer un métier particulier à cette catégorie de participants.

catégories de transgressions morales (par exemple les cas d'inégalité de traitement) mais pas avec d'autres catégories de transgression (en l'occurrence les nuisances gratuites perpétrées contre autrui). Une alternative serait de concéder que la motivation s'applique uniquement au volet positif de la morale (motivation individuelle à faire le bien) et non au volet négatif (motivation à sanctionner l'immoralité). Enfin on pourrait concéder que le caractère motivant est très faible si bien qu'il s'exprime uniquement dans des conditions très particulières ; dans le cas qui nous occupe il se pourrait que les enseignants aient été motivés par des considérations morales parce qu'un de leurs objectifs est de former le caractère moral de leurs élèves. Evidemment, plus les concessions consenties sont importantes, plus il devient urgent de se demander si la prise de position originale vaut la peine d'être maintenue.

Nous avons vu que les résultats de l'étude sur la punition morale mettent en cause la pertinence des théories de type internaliste. En revanche, l'impact de cette étude sur les positions externaliste et situationnaliste est moins évident car le modèle expérimental n'a pas été pensé pour tester l'effet d'un facteur causal externe sur la punition morale. En revanche les résultats obtenus fournissent des indications utiles ainsi que des pistes pour de futures recherches.

L'élément troublant dans les résultats obtenus provient de la propension des enseignants à punir, au contraire des deux autres catégories de participants. La clef de l'explication pourrait bien être un facteur de contexte non moral qui a déclenché un mécanisme motivationnel. Pour commencer, nous pouvons constater que les enseignants ont pour rôle d'éduquer des élèves. Or l'histoire racontée aux participants (« cover story ») met précisément en scène un professeur qui condamne le mauvais comportement d'une de ses élèves. Il mentionne également de manière explicite que le comportement de cette violoniste perturbe le groupe. Sachant cela, au moins trois facteurs non moraux pourraient être à l'origine de la motivation à punir. Le premier facteur pourrait être une heuristique sociale (utile pour la vie en société) du type « choisis l'action qui se conforme aux attentes des membres de ton entourage ». L'influence de cette heuristique sur le comportement humain est d'ailleurs corroborée par un certain nombre d'études⁴⁶. Etant donné que le rôle social des enseignants est

⁴⁶ C. Bicchieri, E. Xiao, « Do the right thing: but only if others do so », in *Journal of Behavioral Decision Making*, 22/2, 2009, pp. 191-208 ; D. B. M. Haun, M. Tomasello, « Conformity to peer pressure in preschool children », in *Child Development*, 82/6, 2011, pp. 1759-67.

d'éduquer les élèves, que l'éducation inclut la mise au pas des individus indisciplinés, et que la procédure de vote a eu lieu dans une salle remplie de futurs enseignants, il est possible que les participants aient simplement appliqué cette heuristique sociale lors du vote. Alternativement, une heuristique sociale faisant référence à l'influence de la source de l'information pourrait être à l'origine du comportement punitif des enseignants. Cette heuristique pourrait être du type « si un de mes pairs respectés loue/condamne une action, j'ajuste mes choix d'action en conséquence ». Ainsi, la motivation ne viendrait pas tellement d'un jugement moral individuel mais plutôt du fait qu'un pair respecté (le professeur de musique) ait indiqué l'action à suivre via un jugement prescriptif. Un troisième mécanisme causal non moral pour expliquer la motivation de punir pourrait être lié à nos capacités de mise en situation. L'idée est que si la situation décrite nous paraît suffisamment familière, nous pouvons nous imaginer pâtir du comportement de la violoniste ; cette mise en situation fournit la motivation d'agir. Ainsi, les enseignants se seraient mis à la place du professeur outré par le comportement de la violoniste et se seraient senti motivés à punir parce que dans une telle situation, un professeur a tout intérêt à prendre les mesures nécessaires pour maîtriser les attitudes antisociales de ses étudiants pour assurer des conditions de travail optimales en classe. Un quatrième mécanisme causal impliqué dans la réaction des enseignants pourrait être le fait que les gens n'aiment pas penser qu'ils font partie d'une institution amoral. Ainsi les enseignants n'aimeraient pas l'idée que le genre d'institution dont ils sont les acteurs clés (une école) puisse récompenser des individus au caractère moral douteux. La liste des explications possibles est ouverte.

VI. Conclusion

Au terme de cette analyse de l'impact théorique d'une étude sur la punition morale, on comprend qu'une seule expérience ne permet pas de déterminer précisément les processus causaux impliqués dans l'action morale. Il en va de même pour la plupart des problématiques traitées au moyen de l'expérimentation. En revanche, les études empiriques ont l'avantage de soumettre les théories à l'épreuve des faits. Ce faisant, elles renforcent ou ébranlent des positions classiques défendues dans la littérature, ou du moins, incitent à en explorer implications (potentiellement indésirables). Pour établir une analogie, les résultats expérimentaux peuvent être représentés par des petites doses d'un élixir qui renforce le système

immunitaire de certains organismes (situation où des résultats confirment une théorie) tout en agissant comme poison contre d'autres organismes (situation où des résultats infirment une théorie). Un organisme sera peu affecté par une seule dose de poison, mais il se mettra à chanceler sous le coup d'une quantité croissante de doses. Au contraire, si l'élixir 'vendu' comme poison tient davantage du placebo (situation où les résultats peuvent être réinterprétés en faveur de la théorie critiquée), ou si l'élixir renforce le système immunitaire, l'organisme se portera d'autant mieux (la théorie acquiert davantage de crédibilité).

Au-delà de sa capacité à soumettre les théories à l'épreuve des faits, la recherche empirique nous fait prendre conscience de la profondeur et de la complexité de la réalité humaine. A l'instar du débat autour de la motivation morale, la pratique de l'expérimentation incite les penseurs de tous horizons à nuancer leurs propos, à raffiner leurs modèles explicatifs et à imaginer de nouvelles explications causales. Enfin, une expérience scientifique en appelle une autre. Les interprétations concurrentes laissées en suspens au terme d'une étude sont autant d'invitations à développer de nouveaux modèles expérimentaux conçus pour faire le tri parmi les possibilités ouvertes. Un engrenage bien pensé d'études successives sur le même modèle est extrêmement utile pour éclairer les eaux profondes de la pensée et de l'action humaine⁴⁷.

Terminons par quelques remarques générales relatives à l'expérimentation utilisée dans un contexte interdisciplinaire. Nous avons vu au travers d'un exemple que l'expérimentation n'est la chasse gardée d'aucune discipline. De fait, elle est utilisée au sein de toutes sortes de sciences (psychologie, économie, anthropologie, philosophie, etc.) et contribue à l'établissement de ponts entre disciplines. Ces ponts sont possibles dans la mesure où les différents chercheurs se posent des questions similaires. Ils ne sont en revanche pas toujours aisés à établir car, malgré les apparences (en général dues à l'utilisation d'une terminologie commune), les chercheurs issus de différents terrains de réflexion ne se posent pas exactement les mêmes questions et comprennent les mêmes termes de manière subtilement différente. Ce phénomène est dû au fait que les développements scientifiques et les débats internes à chaque discipline orientent largement les questions et définitions d'usage au sein des

⁴⁷ Pour un bel exemple du genre, voir : C. D. Batson, *The altruism question: Toward a social psychological answer*, Hillsdale, Lawrence Erlbaum, 1991.

différentes communautés de chercheurs⁴⁸. Ainsi, la voie interdisciplinaire est semée d'embûches ; d'importants efforts de clarification conceptuelle sont généralement nécessaires pour éviter les confusions et évaluer l'impact théorique de données issues d'autres disciplines.

Pour dépasser ces obstacles, la méthode la plus efficace consiste sans doute à stimuler la discussion et les débats entre acteurs de différents horizons, afin de créer un terreau commun. Lorsque cette pratique est pleinement réalisée, un phénomène intéressant peut être observé : les chercheurs impliqués ne parviennent plus à s'identifier avec une science particulière car l'interdisciplinarité accomplie ne distingue plus entre les disciplines qui l'ont fait naître. Ainsi, dans un contexte interdisciplinaire abouti, un qualificatif tel que « philosophie expérimentale » devient vide de sens.

**Mes remerciements à deux lecteurs anonymes et Florian Cova pour leurs suggestions critiques sur des versions antérieures de cet article ainsi qu'à Michel Chapuisat, Danielle Mersch, Fabrice Clément et Colby Tanner pour leur importante contribution à cette recherche sur la motivation à la punition morale.*

⁴⁸ Pour un exemple typique, voir : C. Clavien, *Je t'aide moi non plus: biologique, comportemental ou psychologique, l'altruisme dans tous ses états*, Paris, Vuibert, 2010 ; C. Clavien, M. Chapuisat, « Altruism Across Disciplines: One Word, Multiple Meanings », in *Biology and Philosophy*, 28, 2013, pp. 125–40.