GIUSEPPE UGAZIO

PERSONALIZED PHILANTHROPY

Estimating moral and financial subjective values to explain preferences for philanthropy

Philanthropy is characterized by a tension between promoting moral values aimed at increasing humanity's quality of life and the material cost incurred to achieve said goal. Material values and preferences are well captured by computational models of choice; however, little is known about moral values: can computation models of choice explain moral preferences?

1. INTRODUCTION

Consider one of the most altruistic decisions one could make: donating one's organs to a person in need. For many, decisions like organ donations rely solely on the moral value associated with the duty of helping others (Courtney and Maxwell 2009; Ongley, Nola, and Malti 2014). For others, however, the decision on whether to donate organs or not results from a tradeoff between the moral values and material (e.g. financial) consequences of such decision. In line with this view, some governments recently passed legislations to regulate the sale of organs (Tabarrok 2010). The case of organ donation exemplifies a tension typical of philanthropic decisions: estimating the moral value of an action and comparing this value with the material costs that such action entail. In order to fully characterize philanthropic preferences and behavior, it is therefore crucial to fully understand how a person computes moral and material values, and how these are traded off by decision-makers.

While behavioral and neuroeconomics developed computational models to yield an accurate estimation of individual preferences for material goods (Glimcher and Fehr 2014; Kahneman 2009; Mullainathan and Thaler 2015; Volkman 2007), to date we do not have any strong evidence to inform us on how subjective moral values underlying moral preferences could be estimated. Developing an accurate measure for moral preferences will then allow us to study how moral and material values are compared and traded off by our brains in order to determine philanthropic preferences. In this paper, I propose a theoretical framework for developing computational models capable of accurately capturing subjective moral values. Jointly with the existing models used to estimate subjective values of material goods, these models could be used to estimate individuals' philanthropic preferences, defined as the result of an interaction between moral and material values.

GIUSEPPE UGAZIO, EDMOND DE ROTHSCHILD FOUNDATIONS CHAIRED ASSISTANT PROFESSOR IN BEHAVIORAL PHILANTHROPY GENEVA FINANCE RESEARCH INSTITUTE, UNIVERSITY OF GENEVA, GIUSEPPE.UGAZIO@UNIGE.CH

2. MORAL DECISION-MAKING

Having a clear understanding of how humans make moral decisions is of critical importance on many practical levels. Indeed, moral decisions with severe consequences pervade many aspects of human life: for instance, doctors routinely face such decisions in the context of organ transplantations (e.g., if there are several candidate patients for one organ, [Courtney and Maxwell 2009]). Other examples involve peacekeeping soldiers who may have to decide whether to turn away refugees from an already full refugee camp, or rescuers who may have to decide whether to obey orders and refrain from saving immigrants crossing the Mediterranean sea (Krosch, Figner, and Weber 2012).

Given their importance, one can find a rich literature that looks at many facets of moral decisions. To date, however, the goal of most scientific approaches to morality has been to understand which elements influence human moral decisions. Previous research from several disciplines, including developmental (Kohlberg 1971) and cognitive psychology (Haidt 2012), neurosciences (Joshua David Greene 2015), and economics (Sanfey 2007), has highlighted the importance of several affective and cognitive processes that may contribute to moral decisions, including emotions (Schnall et al. 2008; Ugazio, Lamm, and Singer 2012; Valdesolo and Desteno 2006), estimates of the intentionality (Young and Saxe 2008) and deservingness (Kliemann et al. 2008) of the subject of a moral decision, and calculation of the probability and magnitude (Shenhav and Greene 2010) of the possible consequences of a moral decision.

However, given this extraordinary literature on the neurobiological and psychological mechanisms supporting moral decisions, it is remarkable how little we understand about

how differences in moral preferences across individuals arise. In other words, we know rather well what may influence moral decisions, but we do not know how the underlying preferences guiding these decisions are formed. Given the substantial role that moral preferences play in guiding philanthropic behavior, it is of critical importance to fill this gap of knowledge. One promising way to achieve this is to approach individual differences in moral preferences from the viewpoint of the *value* that people place on each of the moral options considered. Computations of values of choice options play a crucial role in many other forms of decision-making (e.g., economic for a review see [Schultz 2006]) and recent studies have proposed that comparable neural representations of value may also underlie moral decisions (Shenhav and Greene 2010). Thus, it is plausible to expect that we can account for individual differences in moral preferences by estimating how individuals estimate the values of the moral choice options they are considering.

"It is therefore hard to predict to which extent moral and financial subjective value estimations rely on shared psychological decision mechanism and are supported by similar neural regions and functions."

Consider a practical application of such approach: the context of moral dilemmas, such as the "Trolley Dilemma" (Thomson 2008), that require a person to decide if it is morally required of them to directly harm a smaller number of people in order to save a greater one. Here is a brief description of the dilemma: A runaway trolley threatens to kill five people. The only way to save the five people is to push a stranger off the bridge, onto the tracks below. He will die if you do this, but his body will stop the trolley from reaching the others. Is it morally permissible to push this stranger off the bridge in front of the trolley?

Existing studies suggest that, on average, approximately seven people out of ten judge killing the stranger morally forbidden (Greene et al. 2004; Greene 2015), even if this would lead to saving five people. From a purely utilitarian view, this judgment is counterintuitive, as saving the lives of five people should have higher utility than not harming one person. The frequently observed refusal to endorse the harmful action has been suggested to result from a negative emotional reaction towards this action (Greene 2009). But: where do the individual differences between those who consider pushing morally required vs. forbidden come from?

3. ESTIMATING MORAL PREFERENCES: A VALUE-COMPUTATIONAL APPROACH

From a value perspective, the decision whether to sacrifice one human life in order to save a number of others involves computation and comparison of decision values for both moral options. Therefore, from this perspective, individual differences in moral decisions (i.e., whether it is morally required to sacrifice one human life in order to save five others) result directly from such value comparisons: for some, the subjective value of harming one human life exceeds that of saving five, hence they think it is morally inappropriate to sacrifice the one to save the five, while for others, who consider the sacrifice morally required, their subjective value of saving five lives exceeds that of sacrificing one life.

Concretely, it is possible to test if moral preferences can be accurately captured through estimates of subjective values, by parametrically manipulating two of the measurable elements composing moral decisions in the described dilemmas: *Magnitude* (i.e. the number) of the lives that one may save and moral *deservingness* of the people whose lives are at stake (for instance, if a person has been previously convicted for a felony). These two factors would not only be easy to manipulate experimentally, but more importantly have also been proposed to play an important role in moral decision-making (Kliemann et al. 2008; Shenhav and Greene 2010). By recording moral decisions in such dilemmas, it is thus possible to estimate for each respondent the influence of both of these factors on moral choice: the manipulation of magnitude is important for determining the moral value placed on the life that may be sacrificed. This can be achieved by estimating the minimum number of lives required for each participant to switch from the decision against the sacrifice to the choice for the sacrifice. The manipulation of deservingness, on the other hand, is important to estimate how moral values are *discounted* by deservingness (i.e., how decreases in deservingness may result in lower moral values that influence moral judgments).

"The intertemporal choice task is a standard tool used in behavioral economics that requires participants to decide between receiving a smaller amount of money soon and a larger amount of money after a longer period of time."

4. COMPARING DECISION PROCESSES FOR MORAL AND MATERIAL VALUES

Intriguingly, this moral decision framework is deliberately designed to closely match a task frequently used in economics: i.e. the well-known intertemporal choice task (McClure et al. 2007). This task is used to estimate the economic preferences for rewards of material goods over time, as it entails decisions about financial or other primary rewards that can be received at varying times in the future. More in detail, the intertemporal choice task is a standard tool used in behavioral economics that requires participants to decide be-

tween receiving a smaller amount of money soon and a larger amount of money after a longer period of time. Studies with such standard tasks usually vary two variables, a) the difference between the smaller and the larger amount and b) the dimension of the time interval that one has to wait in order to get the larger amount of money (Green et al. 2004; Green and Myerson 2004). Several studies have shown that humans and animals have a strong preference (or impulse) towards rewards delivered sooner, in particular if a reward may be obtained immediately (McClure et al. 2007). This preference has been explained by a mechanism called *temporal discounting*, referring to a de-valuation of the delayed reward with increasing time to the reward receipt, in a non-linear (often hyperbolical) fashion: theories of temporal discounting propose that the impulse of taking the immediate rewards is stronger the more one has to wait to receive the delayed reward.

"Several studies have shown that humans and animals have a strong preference (or impulse) towards rewards delivered sooner, in particular if a reward may be obtained immediately (McClure et al. 2007)."

Relying on the two tasks described, one can therefore measure in the same individuals their moral and financial preferences, respectively. Given the structural similarities between the two tasks, the proposed experimental paradigm provides an optimal framework to identify independently the neuro-cognitive mechanisms involved in economic and moral decisions respectively, as both entail the same fundamental elements involved in value discounting: magnitude-based valuation (the amount of lives saved/reward obtained) and a devaluing/discounting element (deservingness of the person to harm/time to wait in order to receive the reward). More in detail, one can crucially assess which elements of moral and economic decisions rely on shared vs. specific choice mechanisms. Leveraging on the similarity between estimations of discounting of moral and financial values is crucial, as it allows us to test if these two value discounting processes resemble each other – and potentially draw on similar brain processes. Indeed, with the aid of neuroscientific methods, such as functional magnetic resonance imaging (fMRI), one could test where in the brain these processes are computed and, most importantly, if individual differences in these two tasks result from overlapping vs. distinct neural processes.

5. COMPARING THE NEURAL BASIS OF MORAL AND FINANCIAL PREFERENCES

Whether moral and financial values are processed by similar or different psychological and brain processes has not been studied thoroughly to date. In particular, value computations and discounting have rarely been studied in the moral decision-making domain. It is therefore hard to predict to which extent moral and financial subjective value estimations rely on shared psychological decision mechanism and are supported by similar neural regions and functions. One can expect that the two types of subjective values rely on partially similar brain processes: value computations based on magnitude are routinely found to correlate with neural activity in the striatum, the ventral-medial prefrontal cortex (vmPFC) and the medial orbitofrontal cortex (mOFC) across various choice contexts (e.g. concerning purchases (Hare et al. 2010); or financial rewards, (Kuhnen and Knutson 2005). Indeed, one previous study looking at estimations of expected values found similar patterns of neural activity correlating with magnitude also in a moral context (Shenhav and Greene 2010). Similarly, another study found that the subjective value of financial rewards in intertemporal choices correlated with activity with these brain areas (Kable and Glimcher 2007). Thus, this evidence suggests that moral and economic decision tasks may involve similar brain areas (Figner et al. 2010; Kable and Glimcher 2007; Shenhav and Greene 2010), although this overlap has never been directly investigated.

On the other hand, some behavioral studies proposed evidence supporting the view that moral and material values are not treated similarly by the brain, as in the case of crowding out (Frey and Jegen 2001). If this view is correct, and moral value computations are represented by distinct psychological and neural mechanisms, one can expect some differentiation in the neural mechanisms involved in discounting moral and economic values. This perspective finds support also by considering the very different nature of the element driving discounting across the two tasks (i.e. a more social element such as deservingness for moral value, and time for the economic value). Relying on indirect evidence from existing neuroscientific studies of moral decision-making, it is plausible to hypothesize that one of the brain areas poten-

tially specialized in computing moral subjective values, and not financial values, is the right temporo-parietal junction (rTPJ). This brain area has been found to be involved in the representation of social situations and processing attributions of social intentions, theory of mind, and false believes (Cushman and Young 2011; Saxe, Carey, and Kanwisher 2004; Sip et al. 2008). Instead, as mentioned above, with respect to brain areas specifically involved in processing subjective financial values, one can expect to identify neural mechanisms embedded in more prefrontal brain areas, particularly in the vmPFC (Kable & Glimcher 2010).

6. CONCLUSION: ESTIMATING PHILANTHROPIC PREFERENCES

In this paper, philanthropic behavior is proposed to be driven by an interaction of moral and material preferences. To fully understand philanthropic preferences, I argued that we need to be able to a) measure accurately the subjective values underlying moral and material values, and b) compare the decision mechanisms that determine moral and financial preferences, respectively. While the latter have been analyzed in detail, the former need to be studied more in depth. To this end, I proposed a theoretical framework and, more concretely, an experimental paradigm that would allow us to achieve these aims. If this endeavor proves to be successful, one can use this theoretical framework and similar experimental designs to investigate how moral and material preferences complement and compete with each other in order to shape philanthropic preferences. Ultimately, being able to capture philanthropic preferences will allow us to build scientifically informed models that predict under which circumstances people will be more likely engaged in philanthropic activities. ■

"To fully understand philanthropic preferences, I argued that we need to be able to a) measure accurately the subjective values underlying moral and material values, and b) compare the decision mechanisms that determine moral and financial preferences, respectively."

References: ▶ Courtney, Aisling E., and Alexander P. Maxwell. 2009. "The Challenge of Doing What Is Right in Renal Transplantation: Balancing Equity and Utility." Nephron – Clinical Practice 111(1). ▶ Cushman, Fiery, and Liane Young. 2011. "Patterns of Moral Judgment Derive from Nonmoral Psychological Representations." Cognitive Science 35(6): 1052–75. ▶ Figner, Bernd et al. 2010. "Lateral Prefrontal Cortex and Self-Control in Intertemporal Choice." Nature Neuroscience 13(5): 538–39. http://dx.doi.org/10.1038/nn.2516. ▶ Frey, Bruno S., and Reto Jegen. 2001. "Motivational Interactions: Effects on Behaviour." Annales d'Économie et de Statistique. ▶ Glimcher, Paul W., and Ernst Fehr. 2014. "Introduction: A Brief History of Neuroeconomics." Neuroeconomics: xvii–xxviii. http://linkinghub.elsevier.com/retrieve/pii/B9780124160088000358. ▶ Green, Leonard et al. 2004. "Discounting of Delayed Food Rewards in Pigeons and Rats: Is There a Magnitude Effect?" Journal of the experimental analysis of behavior 81(1): 39–50. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1284970&tool=pmcentrez&rendertype=abstract. ▶ Green, Leonard, and Joel Myerson. 2004. "A Discounting Framework for Choice with Delayed and Probabilistic Rewards." Psychological Bulletin 130(5): 769–92. ▶ Greene, Joshua D. et al. 2004. "The Neural Bases of Cognitive Conflict and Control in Moral Judgment." 44: 389–400. ▶ 2009. "The Cognitive Neuroscience of Moral Judgment." The cognitive neurosciences 4: 987–99. ▶ Greene, Joshua David. 2015. Moral Tribes?: Emotion, Reason, and the Gap between Us and Them. ▶ Haidt, J. 2012. "The Righteous Mind." Why Good People are Divided by Politics and Religion ... (January): 1–508. http://www.sce.cornell.edu/sce/altschuler/pdf/altschuler_review_20120301_460.pdf%5Cnpapers3://publication/uuid/AE91E6CE-E7BE-4F35-85A6-1007295862C3. ▶ Hare, Todd a, Colin F. Camerer, Daniel T. Knoepfle, and Antonio Rangel. 2010. "Value Computations in Ventral Medial Prefrontal Cortex during Charitable Decision Making Incorporate Input from Regions Involved in Social Cognition." The Journal of neuroscience?: the official journal of the Society for Neuroscience 30(2): 583–90. ▶ Kable, Joseph W., and Paul W. Glimcher. 2007. "The Neural Correlates of Subjective Value during Intertemporal Choice." Nature neuroscience 10(12): 1625–33. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2845395&tool=pmcentrez&rendertype=abstract. ▶ Kahneman, Daniel. 2009. "Remarks on Neuroeconomics." Neuroeconomics: 523–26. ▶ Kliemann, Dorit, Liane Young, Jonathan Scholz, and Rebecca Saxe. 2008. "The Influence of Prior Record on Moral Judgment." Neuropsychologia 46(12): 2949–57. http://www.ncbi.nlm.nih.gov/pubmed/18606175. ▶ Kohlberg, Lawrence. 1971. "Stages of Moral Development." Moral education: 23–92. ▶ Krosch, AR, B. Figner, and EU Weber. 2012. "Choice Processes and Their Post-Decisional Consequences in Morally Conflicting Decisions." Judgment and Decision Making. ▶ Kuhnen, Camelia M., and Brian Knutson. 2005. "The Neural Basis of Financial Risk Taking." Neuron 47(5): 763–70. ▶ McClure, Samuel M. et al. 2007. "Time Discounting for Primary Rewards." Journal of Neuroscience 27(21): 5796–5804. http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.4246-06.2007. ▶ Mullainathan, Sendhil, and Richard H. Thaler. 2015. "Behavioral Economics." In International Encyclopedia of the Social & Behavioral Sciences: Second Edition. ▶ Ongley, Sophia F., Marta Nola, and Tina Malti. 2014. "Children's Giving: Moral Reasoning and Moral Emotions in the Development of Donation Behaviors." Frontiers in Psychology. ▶ Sanfey, Alan G. 2007. "Decision Neuroscience: New Directions in Studies of Judgment and Decision Making." Current Directions in Psychological Science 16(3): 151–55. ▶ Saxe, R., S. Carey, and N. Kanwisher. 2004. "Understanding Other Minds: Linking Developmental Psychology and Functional Neuroimaging." Annual review of psychology 55: 87–124. http://www.ncbi.nlm.nih.gov/pubmed/14744211. ▶ Schnall, Simone et al. 2008. "Disgust as Embodied Moral Judgment." ▶ Schultz, Wolfram. 2006. "Behavioral Theories and the Neurophysiology of Reward." Annual review of psychology 57: 87–115. http://www.ncbi.nlm.nih.gov/pubmed/16318590. ▶ Shenhav, Amitai, and Joshua D. Greene. 2010. "Moral Judgments Recruit Domain-General Valuation Mechanisms to Integrate Representations of Probability and Magnitude." Neuron 67(4): 667–77. http://linkinghub.elsevier.com/retrieve/pii/S0896627310005817. ▶ Sip, Kamila E., Andreas Roepstorff, William McGregor, and Chris D. Frith. 2008. "Detecting Deception: The Scope and Limits." Trends in Cognitive Sciences 12(2): 48–53. ▶ Tabarrok, Alex. 2010. "The Meat Market." wsj.com. ▶ Thomson, Judith Jarvis. 2008. "Turning the Trolley." Philosophy and Public Affairs 36(4): 359–74. ▶ Ugazio, Giuseppe, Claus Lamm, and Tania Singer. 2012. "The Role of Emotions for Moral Judgments Depends on the Type of Emotion and Moral Scenario." Emotion 12(3): 579–90. ▶ Valdesolo, Piercarlo, and David Desteno. 2006. "Manipulations of Emotional Context Shape Moral Judgment." Psychological Science 17(6): 476–77. ▶ Volkman, Jacqueline M. 2007. "Behavioral Aspects of Financial Decision-Making." Dissertation Abstracts International Section A: Humanities and Social Sciences. ▶ Young, Liane, and James Dungan. 2012. "Where in the Brain Is Morality? Everywhere and Maybe Nowhere." Social Neuroscience 7 (April 2015): 1–10. ▶ Young, Liane, and Rebecca Saxe. 2008. "The Neural Basis of Belief Encoding and Integration in Moral Judgment." NeuroImage 40(4): 1912–20.