

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/111715>

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

The social character of moral reasoning

Nick Chater, Hossam Zeitoun & Tigran Melkonyan

Warwick Business School,

University of Warwick.

May provides a compelling case that reasoning is central to moral psychology. In practice, many morally significant decisions involve several moral agents whose actions are interdependent; and agents embedded in society. We suggest that social life, and the rich patterns of reasoning that underpin it, are ethical through and through.

Word count:

Abstract: 50

Main text: 2007

May makes a compelling case for the importance of moral reasoning that inform our ethical judgments and actions. This conclusion is reinforced if we widen our scope to consider situations in which morality seems to depend on not only our own actions, but also the actions of others; and, more broadly, ethics concerns rules and policies for the smooth operation of society, in which each person has specific roles and responsibilities. Moral agents are not lone and omnipotent decision-makers, setting the course of a moral microcosm in which they have jurisdiction (e.g., whether to pull the lever in a trolley problem; whom to rescue in a shipwreck; and so on). They are instead active participants, alongside other active participants, in an endlessly complex social world of families, organizations, nations, professions, customs, conventions, norms and laws.

Consider, for example, the well-known transplant dilemma (Thomson, 1985) that May discusses in Chapter 3. The dilemma is whether a surgeon should forcibly remove the organs of one person to save the lives of five others, and hence apparently generate a net gain, from a utilitarian point of view (note that such actions are not allowed by the Pareto criterion in welfare economics). The extreme concern that most of us feel about this action might, of course, be set aside as emotionally-driven squeamishness. But, on reflection, our distaste surely has a credible basis in moral reasoning. A world in which such practices were sanctioned would be one in which patients would refuse to go to hospital, staff would flee for their lives, doctors would be feared rather than welcomed, and surgeons would resign en masse. To sanction such behavior would be to risk pulling apart the entire fabric of the healthcare system, and to rip up fundamental tenets of law and policing. Indeed, an enthusiastic advocate of the utilitarian approach might attempt to prosecute doctors for refusing to make such transplants (leading, by assumption, to a net “loss” of four lives); and to prosecute police, prison officers and judiciary who refuse to comply. Such considerations seem to provide ample reason to explain our revulsion. Indeed, these considerations would surely be in the forefront of the minds of physicians, medical ethicists, and government policy makers, were the possibility of allowing such transplants a politically live issue. (May rightly makes a related point in terms of the reasons people give—regarding guilt, long-term psychological harm, shame or, potentially, undermining of religious beliefs and practices—when justifying ‘harmless’ taboo violations (see Royzman, Kim, & Leeman, 2015)).

Some moral philosophers and moral psychologists might wish to wave aside such concerns, insisting that we focus only on the microcosm of the ‘thought experiment,’ and nothing beyond it (as if, for the purposes of the example, the world consisted of six patients, a surgeon, and nothing more; or of an isolated careening trolley car, some people it may strike, some levers, and one or two hapless bystanders). But this *asocial* idealization, in which the ethical dilemma is disconnected from wider society, will be fundamentally misguided if, as we suggest, the fundamental rationale for our ethical principles and intuitions is the well-functioning of that society. Indeed, attempting to introspect, or collect data, on such putatively isolated moral problems may be akin to attempting to understand shoaling behavior by studying the movements of an isolated fish, out of water.

Indeed, such isolated examples are inevitably likely to yield limited insight into the rich web of moral reasoning which guides social life, because they are deliberately disconnected from that web. A parallel tack in epistemology would yield similar conclusions: suppose people were asked what could be concluded *solely* from finding that the light passing through a prism forms a spectrum, or that feathers and cannon-balls fall at the same speed in a vacuum. If such questions must be answered without any connection to the rest of our knowledge of the physical world, then few conclusions will

be forthcoming; and one might be tempted to conclude that reason plays little role in science too. But, again, the disembodied example is stripped of useful reasoning—because the practically-relevant reasoning concerns the relationship between specific experiments (or moral dilemmas) and the web of knowledge in which they are embedded.

Note, too, that the richness and complexity of moral reasons depends on our ‘location’ in the social world—a matter ignored in many philosophical examples and psychological experiments. Consider, for example, the moral dilemma faced by a college-admissions tutor, who realizes that an applicant is the daughter of a close friend. The applicant’s test scores are just below the cut-off; but the tutor knows that the daughter has a phobia of tests and performs much worse than she could. For most of us, the case seems clear-cut: the tutor should apply the same rules to everyone or, and probably preferably, refer this student to a colleague. Why? Because there is an agreed process for impartially handling applications; and the admissions tutor’s role is to follow that process. These are the reasons that the tutor would presumably provide to explain making no exception. The consequences for the applicant (and for the applicant whom she might displace) are not relevant considerations (conversely, were the tutor to make an exception, a great deal of reasoning would be provided—the extremity of the case, the potential loss of a shining academic star, the personal devastation, and so on).

The moral psychologist or moral philosopher might be tempted to respond: but these reasons are all about why behaving in a particular way discharges a person’s job—here, what is *right for an admissions tutor*. But perhaps morality is about what is right *simpliciter*. We suggest that this type of response goes to the heart of the problem. If moral reasoning guides social behavior and the roles and responsibilities each of us has in society, then the very idea of ‘right’—independent of roles and responsibilities—verges on incoherence. The moral decision makers are not distant and omnipotent decision makers; they are real human beings, struggling with their conflicting roles of, here, being admissions tutor and helpful family friend.

As noted, much work in moral philosophy and moral psychology is not merely asocial and concerned with decision makers with no ‘location’ in the social setting. Much such work appears, moreover, to be directed at a hypothetical omnipotent decision maker, rather than at participants with specific roles in an unfolding drama (see Sugden, 2018, for a closely related argument in economics).

Often, the question at the heart of ethical debate—and implicit in many related psychological studies—is close to: what would you decide should be done here *if you ruled the world* (benevolently, of course)? But this is surely an unhelpful viewpoint! Each of us makes our ethical decisions locked within not just a specific role, with limited power, but at the mercy of many other decision makers, each making their own ethical decisions. And, worse, the results of our choices are interdependent, in potentially complex ways. Thus, we might expect that a good part of ethical reasoning will concern how we coordinate and negotiate our way through a mass of other people, each coordinating and negotiating as we are. And then the goal of ethics might properly be directed to helping individual citizens manage such challenges from their specific vantage point.

Consider, for example, a variation of the much-discussed trolley car example, originated with Foot (1967). Suppose that the trolley is hurtling towards 10 people whom it will kill instantly. A set of 5 people each has independent access to a switch that will divert the trolley to a parallel track. Unfortunately, this switch works on a toggle: each time the switch is pressed, the train flips track

again. So, if an odd number of switches are pressed, then disaster will be averted; if an even number is pressed, it is not.

Imagine, to start with, that it is common ground that all five people are well intentioned: they want to avoid calamity. But, still, what is the right thing to do?

Suppose, for example, that A knew that B, C, D and E will do nothing. Then A should, of course, press the switch. But perhaps one of the others will press the switch; then A doing the same will cause, rather than prevent, disaster. Or perhaps two of the others will press the switch; in which case A must press, too. And the others, B, C, D, and E, face the same dilemma of course.

Note, though, that there is an intuitively elegant solution to this puzzle, which will doubtless already have occurred to the reader. Since there is an odd number of players, if all five people press the switch, then success is guaranteed.

Suppose, that each person notices this, each therefore presses the switch, and the good outcome is obtained. The reasoning involved here is rather subtle. One way to reconstruct this reasoning is for each player to ask themselves: if we could communicate, what policy would we agree? If it is 'obvious' that the simplest and most general policy is that everyone chooses to switch, and that they would agree this policy were they able to communicate, then communication is unnecessary. A, B, C, D and E simply imagine the outcome of the hypothetical process of reaching an agreement and implement the result. This is the type of reasoning we call virtual bargaining (Melkonyan, Zeitoun, & Chater, 2018; Misyak, Melkonyan, Zeitoun, & Chater, 2014)—people imagine the outcome of a hypothetical bargaining process; and directly implement the agreement.

Notice, crucially, from a virtual bargaining standpoint, ethical theory focuses on advising individuals about what they should do, given their collective challenge; it helps people align their behaviors to jointly achieve a successful outcome. The fundamental challenge for the moral philosopher is not: what should I command that these people do, if I ran the world? but rather, how might I help advise individuals in this situation to help them collectively bring about a good outcome?

Let us imagine, for a moment, that E chooses not to press the switch, and disaster occurs. What is the moral status of E's action? The others may turn on E and blame her for the disaster: the moral emotions will be dialed up to maximum. But notice that reasoning is the source. Suppose E tried the following retort: "Well, if any one of us had done something different, all would have been well. I'm not especially to blame" (and indeed, many models of responsibility, e.g., Chockler & Halpern, 2004, have difficulty with this type of case). This would be met with utmost scorn. But suppose E turned out to be misinformed—unlike the others, E had been told nothing about the functioning of the button; or perhaps E had been told there were six, not five, people with buttons. Then E is absolved of guilt; our collective rage might be directed at F, who deliberately, and with malice aforethought, misled E to bring out disaster.

Our moral emotions are directed at who seems to be to blame; and who seems to be to blame (no one, E, or F) depends on the outcome of subtle moral reasoning about hypothetical agreements.

A final possible objection. Can the proponent of an emotion-based account of moral psychology suggest that all this reasoning is not *moral*, but is simply reasoning about goal-directed social behavior (and that the goal in this case is saving lives, which is where morality enters)? We propose

the very opposite: that morality suffuses every aspect of social behavior; that the prescriptions of what we should and should not do, which rules we should live by, what is worthy of praise and blame, are moral through and through. Moral reasoning is the foundation for society in much the way that reasoning about the external world is the foundation for science. Laws, money, institutions, roles, rights, responsibilities and governments are all products of moral reasoning. May is right: moral reasoning is of primary importance. Indeed, the creation, critique and defense of moral reasons, large and small, is the essence of our emotional, social and political lives.

References

- Chockler, H., & Halpern, J. Y. 2004. Responsibility and blame: A structural-model approach. *Journal of Artificial Intelligence Research*, 22: 93-115.
- Foot, P. 1967. The problem of abortion and the doctrine of the double effect. *Oxford Review*, 5: 5-15.
- Melkonyan, T., Zeitoun, H., & Chater, N. 2018. Collusion in Bertrand versus Cournot competition: A virtual bargaining approach. *Management Science*: forthcoming; available on: <https://pubsonline.informs.org/doi/full/10.1287/mnsc.2017.2878>.
- Misyak, J. B., Melkonyan, T., Zeitoun, H., & Chater, N. 2014. Unwritten rules: Virtual bargaining underpins social interaction, culture, and society. *Trends in Cognitive Sciences*, 18(10): 512-519.
- Royzman, E. B., Kim, K., & Leeman, R. F. 2015. The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect. *Judgment and Decision Making*, 10(4): 296-313.
- Sugden, R. 2018. *The community of advantage: A behavioural economist's defence of the market*. Oxford: Oxford University Press.
- Thomson, J. J. 1985. The trolley problem. *Yale Law Journal*, 94(6): 1395-1415.