

The Cognitive Mechanisms of Contractualist Moral Decision-Making

Sydney Levine^{1,2} (smlevine@mit.edu), Max Kleiman-Weiner¹ (maxkw@mit.edu),

Nichols Chater³ (nick.chater@wbs.ac.uk), Fiery Cushman² (cushman@fas.harvard.edu) & Joshua B. Tenenbaum¹ (jbt@mit.edu)

¹Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139

²Psychology, Harvard University, Cambridge MA, 02138

³Business School, University of Warwick, Coventry, UK CV4 7AL

Abstract

Contractualism is a theory of moral philosophy that posits that an act is morally permissible if all the parties relevantly affected by the act could reasonably agree to it. We take this theory of moral philosophy as an inspiration for a theory of moral cognition. In this paper, we present evidence that subjects have contractualist intuitions and use explicit contractualist reasoning. These data are poorly accounted for by current theories of moral cognition which rely mostly on the use of rules or calculations of consequences. We sketch out a rational model that captures these phenomena by predicting subjects moral judgments as a function of their representation of the interests of agents who are engaged in a mentally simulated bargaining process. We conclude by discussing how a computational cognitive science of contractualism fits into a utility-based unified theory of moral cognition, which integrates elements of rule-based, consequence-based and contract-based cognitive mechanisms.

Keywords: moral judgment; social cognition; contractualism

Introduction

Philosophical moral theories offer a window into ordinary moral thinking. Indeed, moral psychology has made great strides recently by considering how our minds might mirror the distinction between two particular moral theories: consequentialism and deontology (J. D. Greene, 2008). Crudely characterized, consequentialism posits that moral permissibility is determined solely by the consequences of an act. Deontology posits that it is determined instead by conformity to rules, which are often stated in terms of duties, prohibitions and rights. Our most promising moral theories describe the mechanisms our minds use to make moral judgments using either psychological rules (Mikhail, 2007; Nichols, 2004), calculations of the utilities of outcomes (Harsanyi, 1978), or both (J. Greene, 2014; Cushman, 2013; Crockett, 2013; Kleiman-Weiner, Gerstenberg, Levine, & Tenenbaum, 2015).

Curiously, a third major family of philosophical moral theories—contractualism—is mostly neglected in contemporary moral psychology (for a notable exception, see (Baumard, 2016)). Contractualism holds that an act is morally permissible if all the parties relevantly affected by the act could reasonably agree to it (or at least not reasonably reject it). Put simply, it revolves around a kind of idealized bargaining.

Here, across several studies, we present evidence for “intuitive contractualism”—i.e., a dimension of our moral psychology that mirrors key features of contractualist moral philosophy.

What is contractualism? Like consequentialism and deontology, there are various forms of contractualism which sig-

nificantly differ on their details (e.g., (Scanlon, 1998; Habermas, 1990; Rehg, 1994; Parfit, 2011); for contractualism as a way of setting up a political system, see (Rawls, 2009). What is central and common to all the views, however, is that the moral acceptability of an action is determined by taking into account the interests of all the affected parties. Often this is modeled as the outcome of a fully rational dialog that would take place between the affected parties. A fair conclusion would be reached by way of the “unforced force of the better argument” in Habermas’s famous words (Habermas, 1996). In practice, of course, affected parties rarely participate in fully rational dialog. Contractualists hold that we should simulate it in order to determine how to treat each other.

Intuitive contractualism as virtual bargaining We take this ideal as inspiration for a psychological model of moral decision-making that uses a *virtual* bargaining process. Virtual bargaining (c.f. (Misyak, Melkonyan, Zeitoun, & Chater, 2014)) involves running a mental simulation of how a conversation would go between all the affected parties.

One attractive feature of this model is that it implies situational flexibility.¹ Indeed, we often seem to treat rules as principles that help guide our behavior and make social interactions possible, but that can be renegotiated if everyone who would be affected would agree to another arrangement. For example, it is typically prohibited for someone to cause bodily harm to someone else, but if pushing someone out of the way of a car would save their life, it may actually be obligatory. From a contractualist perspective, all the parties affected by the rule would agree to the temporary suspension of the rule to deal with the critical situation that arose.

Such flexibility is useful because it accommodates unusual cases, which may fall beyond the scope of rule-based approaches. Such unusual cases will arise in a changing world: New objects, new kinds of causal mechanisms, new kinds of technologies and new possibilities for social interaction are unavoidable. Moreover, strangers become friends, friends become enemies, and what is in one moment virtuous we can later be seen as a vice.

¹In fact, even legal systems that see their rules as literally set in stone, like Judaism’s view of the ten commandments, make room for putting aside those rules in dire circumstances. The Talmud, an extensive commentary on the Bible which is the foundation for much of Jewish law, explains that a verse in Psalms (“Make void Your Torah; it is the time to work for the Lord”, Psalms 119:126) should be interpreted to mean that “occasionally it is necessary to negate biblical precepts in order to bolster the Torah” (Berachot 54a).

Experimental approach Employing a case that is quite unusual indeed, we provide evidence for intuitive contractualism. Participants judgments show patterns characteristic of contractualist theories (as opposed to rule-based or consequence-based theories) (Experiment 1) and they use explicitly contractualist reasoning in the form of mentally simulated bargains between the affected parties (Experiment 2). Finally, we begin to develop a computational account of virtual bargaining and describe its quantitative fit to data (Experiment 3). We conclude by considering how intuitive contractualism may be integrated into a unified approach to moral psychology including the consideration of rules, consequences, and contracts in a utility-based framework.

Experiment 1

The first study was designed to provide initial evidence that subjects have intuitions about moral acceptability that can be broadly described as contractualist, which are not captured by strong versions of consequentialism or deontology.² The goal of the study was to show that subjects (1) spontaneously determine who the “affected parties” are, and then (2) favor a course of action that takes the interests of those affected parties into account, as if by bargaining. These features are essential to the contractualist approach.

Materials and Methods

A version of the following vignette was shown to each subject ($n=151$ MTURK participants): “A mysterious but very wealthy stranger arrives at Hank’s doorstep with a strange request. He asks Hank whether he would be willing to paint his next-door neighbors house blue. If Hank agrees, he will receive a sum of \$1 million which Hank can use however he likes. Otherwise, the stranger will leave and everything remains as before. Rather inconveniently, Hank’s neighbor is away on vacation, and cannot be communicated with for the next week—but the mysterious stranger requires an answer today. The man shows the \$1 million dollars in cash to Hank right then and there. Hank can either take the \$1 million and paint the house, without permission from his neighbor, or he can turn the stranger away.”

Following standard contractualist theories, we predicted that participants would count the neighbor as an “affected party” in this case (hereafter, the Neighbor’s House condition); we therefore contrasted it with two alternative cases for which we predicted the neighbor would not count as affected – for different reasons. In the Own House condition, the stranger requests that Hank paint his own house blue (not his neighbor’s) in order to receive the \$1 million. In this case, the neighbor is clearly not an affected party. In the Not Involved condition, Hank’s neighbor’s house accidentally gets

²Consequentialism and deontology can each be modified in such a way to be able to capture the intuition about the cases we describe here. This itself is an interesting phenomenon that we think may be a useful approach to understanding how to construct a unified theory of moral judgment. We put that point aside for now, and instead focus on “core” versions of each of the theories.

painted blue due to a miscommunication not related to Hank. Incidentally, a stranger appears and gives \$1 million to Hank with no strings attached. In this case also, the neighbor is not an affected party. Despite the fact that the same harm is done to him as in the Neighbor’s House condition, that harm is not related to a moral decision Hank makes.

In all conditions, subjects were asked which of Hank’s options is most morally acceptable. They chose an answer from the following options (order of choices was randomized; bolded titles were not presented):

- **Split:** Accept the offer and give some of the money to the neighbor
- **Donate:** Accept the offer and donate the money to charity
- **Refuse:** Refuse the offer and don’t touch the neighbors house
- **Keep:** Accept the offer and keep all the money

If a subject responded that Hank should give some money to the neighbor, they were asked to indicate (by entering a number from 1-1,000,000) how much money Hank should give the neighbor (which we will refer to henceforth as “side-payment” judgments). For the Neighbor’s House condition, this possibility was designed to be favored by a contractualist approach. In contrast, “Donate” was designed to be favored by a consequentialist approach because it maximizes aggregate welfare, “Refuse” was designed to be favored by a deontological approach because it respects the property rights of the neighbor, and “Keep” was designed to be disfavored by all approaches.

Results There are two main findings from this study. First, in the Neighbor’s House condition, subjects were biased in favor of offering a side-payment to the neighbor (“Split”) as the most morally acceptable response ($\chi^2(3, n=48)=31.3, p < .0001$). Subjects find this option more morally acceptable than “Donate” (i.e., consequentialism) and “Refuse” (i.e., deontology) (see Figure 2). Many subjects favored an even split of the money, which may reflect a fairness concern, and many others favored \$5,000, which may reflect the cost of repainting (i.e., financial compensation; Figure 1).

The second major finding is that in the Own House condition—when the neighbor is not at all involved in the deal between Hank and the stranger—subjects’ patterns of judgments shift dramatically. Subjects in this condition were biased towards responding that it is morally acceptable for Hank to keep the money, with very few subjects responding that Hank should give some to his neighbor ($\chi^2(3, n=53)=14.2, p < .005$). This suggests that the neighbor’s interests should be taken into account only when he is directly impacted by the outcome under consideration (Figure 1).³ The Not Involved condition highlights the fact that it is the neighbor’s involvement in the deal (it is Hank’s painting of his house without his permission) that makes him an affected party, not simply that something bad happened to him

³Of course, who counts as being affected may not always be so clear. We put aside this issue for now.

around the time and place that Hank receives an offer from the stranger. Interestingly, in the Not Involved condition, subjects are most likely to say that the most morally acceptable thing for Hank to do is to give some of the money to his neighbor, just like in the Neighbor’s House condition (no significant difference between the proportion of subjects in each condition that respond that Hank should split the money: $\chi^2(1, n=98) = .33, p = n.s.$). However, the distribution of side-payments (Figure 2) reveals that the vast majority of subjects in the Not Involved condition give a small amount to the neighbor (enough to have his house re-painted in the face of this unfortunate accident) and there is almost no evidence of subjects expressing that it would be most fair to split the money with his neighbor, as they do in the Neighbor’s House condition.

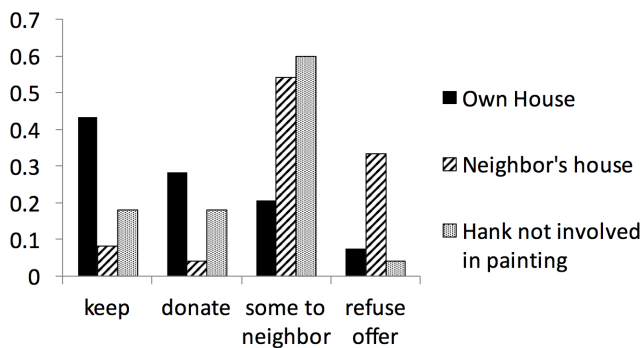


Figure 1: Subjects responded to the question: “Which of those options is most morally acceptable?” Percent of subjects choosing each option is graphed.

Experiment 2

Experiment 1 was designed to provide initial evidence that subjects have moral intuitions that cannot be explained by strong versions of consequentialism and deontology. We suggest that the intuitions revealed in Experiment 1 are instead best captured by a contractualist framework that involves a mental accounting of the interests of the affected parties. Experiment 2 was designed to investigate the nature of this mental accounting. Following a standard contractualist approach, we hypothesized that subjects thought about what each of the affected parties would say if they could speak with one another in a fully robust dialog (e.g., (Habermas, 1990))—in essence, striking a “virtual bargain” with their neighbor.

Materials and Methods Subjects (n=24 MTURK participants) read the Neighbor’s House vignette described in Experiment 1. Instead of being presented with options for Hank’s decision, subjects were asked answer a series of ordered free-response questions about what Hank should do:

1. What should Hank do and why?
2. Is there an ethical way for Hank to accept the offer? Why, or why not?
3. Hank imagines calling his neighbour and imagines what they would talk about if they could. In this imagined con-

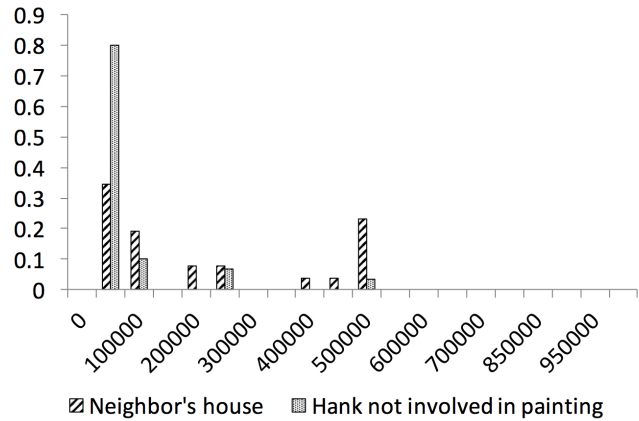


Figure 2: Subjects responded to the question: “How much of the \$1 million should Hank give to his neighbor? (Please enter a number from 1-1,000,000.)” Only subjects who responded that Hank should give some money to the neighbor gave responses to this question. Subjects responses were binned and percent of subjects giving a side-payment judgment in each bin is graphed. No side-payment judgments from the Own House condition are graphed due to the small number of subjects who responded that Hank should give some of the money to the neighbor, as shown in Figure 2.

versation, what conclusion does Hank think he and his neighbour would reach? How does this impact what he thinks he should do?

4. Imagine you are Hank’s neighbor away on vacation. When you come back and find that your house is painted blue, would you be understanding?
5. What (if anything) would you expect Hank to do on your return in order to feel he acted ethically?
6. When you return to find your house painted blue, Hank is apologetic. He explains the story with the mysterious stranger, and says “I felt the right thing to do was to split the money 50-50. I tried to imagine how you’d feel about this, and thought you’d probably think this was a good plan - I hope I got that right.” How would you feel about Hank’s behavior?

The series of questions were designed to reveal elements of virtual bargaining. Each question was coded by two independent coders for the presence or absence of particular features of interest in responses to each question. Disagreements between the coders about the presence of the feature were settled by discussion. The proportion of subjects expressing each feature in their response is reported in Table 1.

Results Most subjects (74%) say that Hank should accept the stranger’s offer (Question 1). The main finding is that, of these, 61% of subjects spontaneously mention side-payments (which were not mentioned in the experiment). We conjecture that this reflects the outcome of “virtual bargaining” with the neighbor. Consistent with this possibility, most subjects say

Table 1: Free Response Answers. The proportion of subjects mentioning each feature of interest was coded for each question.

	Feature coded	
Q1	Accept stranger’s offer	0.74
Q1	Spontaneous mention of side-payments	0.61
Q2	Ethical way to accept	0.43
Q3	Agree after conversation	0.76
Q4	Would understand	0.76
Q5	Would request side-payment	0.78
Q6	Would accept 50/50 split	0.96
All Qs	Any mention of side-payments	0.87
All Qs	Any mention of charity	0.0
All Qs	Any mention of property rights	0.43

that if they were the neighbor who returned to find their house painted blue, they would understand and support Hank’s decision (76%, Q4) while also expecting to be given some of the money (78%, Q5), with almost all subjects saying that they would find a 50/50 split of the money to be an acceptable outcome of this scenario (96%, Q6).

Despite this, only 43% of subjects are willing to say that Hank acts ethically when he accepts the offer (Q2). Perhaps our subject’s use of the word “ethical” may be tied to notions of obeying clear moral rules in the deontological sense, even though many of those same subjects seem to express that those rules (essentially a standing contract between all members of society to respect other people’s property rights) could be usefully renegotiated online to yield an agreement that everyone would find more acceptable than the result that would come about if all the rules were followed.

As an illustration of this phenomenon, one subject who responded that there is no ethical way to accept the offer said, “It’s not ethical but who cares it’s a million dollars you’re neighbor would make fun of you if you didn’t agree to that.” This suggests that there is an understanding between Hank and the neighbor of what the rules of ethics state (don’t violate property rights) but that they would both agree that the obviously correct thing to do would be to accept the offer and, from there, work things out between the two affected parties.

Finally, subjects who said that it was ethical to paint the house sometimes drew on an imagined conversation that the two would have later, for example: “If Hank simply paints the house, once the neighbor gets back, Hank can easily explain his decision to him and he’ll understand totally. Thus, no one was harmed and they both will be richer for it!”

Experiment 3

Experiment 1 suggested that subjects have contractualist intuitions in the blue house dilemma. Experiment 2 provided evidence that those intuitions reflect spontaneous reasoning about how to account for the interests of all the affected parties by conducting a virtual bargaining process involving set-

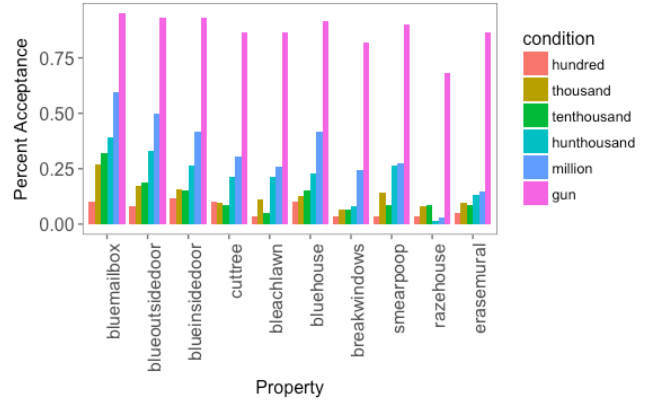


Figure 3: Percent of subjects who judged that agreeing to the offer would be morally acceptable.

ting on side-payments. Experiment 3 complements this finding in a more quantitative manner. Specifically, we assess whether participants’ moral acceptability judgments are predicted by both the stranger’s offer and the compensation they think someone would request to have a harm voluntarily done to them.

Materials and Methods Experiment 3 adapts the basic structure of the Neighbor’s House’ case. In Experiment 3, however, two parameters of the dilemma were manipulated: 1) the amount of money the stranger was offering and 2) the property damage that the stranger requests Hank to carry out in exchange for the money. The offers were \$100; \$1,000; \$10,000; \$100,000; \$1,000,000, or a threat against the life of Hank’s son (“gun” henceforth). The property damages were: painting the neighbor’s mailbox blue, painting the outside of the neighbor’s front door blue, painting the inside of the neighbor’s front door blue, painting the neighbor’s house blue, cutting down a tree in the neighbor’s yard, breaking all the windows in the neighbor’s house, spilling several gallons of bleach on the neighbor’s lawn, smearing dog poop on the neighbor’s front steps, painting over a mural created by neighbor’s daughter, or entirely demolishing the neighbor’s house.

One group of subjects (n=360, MTURK participants, 60 subjects in each condition) was exposed to series of cases each characterized by an offer amount and a property damage type, and were then asked whether it would be morally acceptable for Hank to accept the stranger’s offer. Next, they were asked to imagine that Hank did accept the offer and that the stranger followed through on his promise and delivered the money. Subjects were then asked how much money Hank would owe the neighbor. Each participant judged 10 cases: each level of property damage crossed with a single offer amount. Offer amount was varied between subjects.

A separate group of subjects (n=100, MTURK participants) made judgments about how much compensation they would have to be offered to have the property damages done to them (“compensation demanded” condition).

Results Figure 3 shows subject’s acceptability judgments for each offer/property pair. There are two main trends to be accounted for: (1) as offer goes up, the probability of the offer being judged morally acceptable goes up, (2) as harm goes up (as determined by subjects’ compensation demands), the chance of the offer being morally acceptable goes down.

We hypothesize that, in making their moral acceptability judgments, subjects use the relationship between 1) their representation of the compensation a person would demand to voluntarily agree to some particular harm (times a parameter β) and 2) the amount of money on offer. If there is enough money on offer from the stranger to compensate the neighbor for the harm (plus provide enough money for the trouble, potential emotional and relationship damage and so forth, which is incorporated into the β parameter), the subject can say that accepting the offer would be permissible with the assumption that Hank would then provide some side-payment to the neighbor. We therefore model probability of acceptance (p , in log odds) as a function of offer and amount of compensation demanded:

$$p = \frac{1}{1 + e^{-\gamma(\text{offer} - \beta \text{comp})}}$$

As part of this semi-rational model, we have already explained that we model subjects’ acceptability judgments as involving their representation of someone else’s mind, here, the amount of money that the neighbor would want to be compensated to voluntarily agree to have a harm done to them. What value for compensation, more specifically, is operative in subjects’ decision-making about the moral acceptability of accepting the offer?

Figure 4, which graphs of the distributions of subjects’ compensation demands for each property, shows that there is a wide range of values that someone might demand when faced with the prospect of a particular harm. We assume that our subjects have a representation of this distribution – that is, that they represent that there is a range of compensation demands that someone might make. We model the value for compensation that subjects are using in their moral acceptability judgments as the 90th percentile of the compensation distributions. Put another way, we assume that subjects choose a value for compensation that they are highly confident will be greater than the amount someone would demand in compensation for that particular harm when making moral acceptability judgments.

One additional feature of subjects’ moral acceptability judgments stands out (Figure 3). As the offer from the stranger increases, there seems to be a cap on the amount that that can impact moral acceptability (around 60%), which only jumps to greater than 90% acceptability for the gun condition (when Hank’s son’s life is at stake). This can be expressed by transforming the offer amount into utilities using a loss aversion function (c.f. (Kahneman, 1979)). We like-

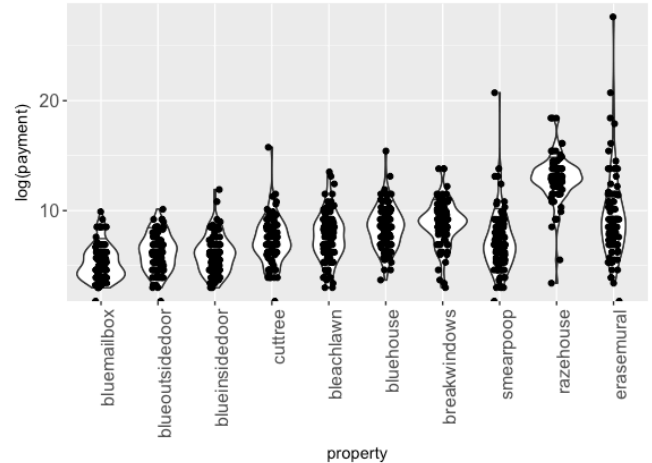


Figure 4: Distributions of subject responses for the compensation they would demand to agree to have some property damage done to them.

wise transformed compensation values into utility space.

$$u(x) = \begin{cases} x^\alpha & \text{if } x \geq 0 \\ -\lambda(-x)^\alpha & \text{if } x < 0 \end{cases}$$

The gun condition, which presents the prospect of the loss of life of a family member, poses the possibility for a utility loss far above the scale of the greatest possible gains or losses in our experiment, which deal almost exclusively with material objects (with the possible exception of the mural case, which may involve sentimental/emotional harm). We arbitrarily set this value to the utility equivalent of \$20 million, but the precise subjective value is not determinable from our data nor critical to our analysis. We fit the α parameter of the model to our data ($\alpha = .25$). (The λ parameter was fixed at 2.25, the value reported by Kahneman & Tversky (Kahneman, 1979)).

Both terms in the model are statistically significant ($p < .0001$) and the model captured most of the variance in the data well: $R = 0.935$ which can be seen in Figure 5.

Conclusion

In three experiments we find that (1) subjects have intuitions that are better explained by appeal to contractualist thinking than consequentialist or deontological thinking and (2) they explicitly use contractualist-style reasoning, in the form of virtual bargaining, to judge moral acceptability. We also defined a simplified rational model of contractualist moral judgment that reasons about the preferences of affected parties, and then uses these preferences to simulate the outcomes of bargaining.

Although intuitive contractualism offers the best account of the specific cases considered here, we do not argue that all moral thought is contractualist. To the contrary, consequentialist and deontological patterns of thought are clearly evident in other cases (J. D. Greene, 2008). Inspired by Derek

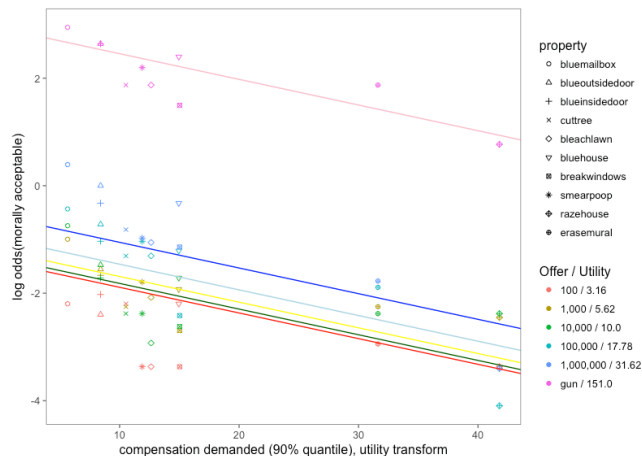


Figure 5: Log odds of the probability of an offer being accepted plotted as a function of compensation demanded (in utility units). Regression lines are drawn through each of offer transformed into utilities. Utility of the offer is treated as continuous in the model.

Parfit’s concept of a “Triple Theory” (Parfit, 2011), we suggest that a unified account of moral thought will encompassing elements of each of these theories.

What might a “Psychological Triple Theory” look like? One possibility is that at the level of the computational theory (Marr Level 1; (Marr, 1982)), the function of moral acceptability judgments is to maximize utility (either at an individual or group level). Put another way, reverse-engineering our moral capacity leads us to conclude that, at its core, the task that the moral organ sets out to do is to optimize utility, broadly construed, under the constraints present in our environment. (Of course, what what features of the world count as having positive utility—and how much utility they have—is an open question. This is the philosophical question of “axiology”, of figuring out which outcomes are actually the best outcomes.) In order to do that, the mind may use different algorithms in different situations. Sometimes, rules are the appropriate way to make moral judgments because they can act as utility-maximizing heuristics (this is essentially the argument behind “rule utilitarianism”.) At other times, rules are insufficiently flexible and the rules should be renegotiated online between all the affected parties to reach an agreement that is best for all concerned. At other times, when the information is particularly clear and available, simply calculating the utilities of outcomes may be the best approach.

As we start to think more seriously about how to program robots who lives with us and are agents in our moral world, rules and utilities may be useful in enabling them to navigate morally-charged situations in line with human ethics, but for circumstances that come up that fall outside the scope of those rules and calculations of consequences, having the AI be able to think about the changing interests of the other agents in the environment may be critical.

In attempting to build a computational cognitive science of contractualism, a range of questions arise. Do we represent the interests of the affected parties as if those people were fully rational, mostly reasonably, or actually realistic (and potentially highly biased) agents? In multi-agent problems, when some agents have directly conflicting interests and some have the same interests, how are the interests of the parties taken into account and weighed against each other? How do we determine who is an affected party at all? When does contractualist reasoning engage and when can we safely rely on rules and/or consequences? All of these provide a rich field for further investigation.

References

- Baumard, N. (2016). *The origins of fairness: How evolution explains our moral nature*. Oxford University Press.
- Crockett, M. J. (2013). Models of morality. *Trends in cognitive sciences*, 17(8), 363–366.
- Cushman, F. (2013). Action, outcome, and value: A dual-system framework for morality. *Personality and social psychology review*, 17(3), 273–292.
- Greene, J. (2014). *Moral tribes: Emotion, reason, and the gap between us and them*. Penguin.
- Greene, J. D. (2008). The secret joke of kants soul. *Moral psychology*, 3, 35–79.
- Habermas, J. (1990). *Moral consciousness and communicative action*. MIT press.
- Habermas, J. (1996). Between facts and norms, trans. william reh. *Polity, Oxford*, 274–328.
- Harsanyi, J. C. (1978). Bayesian decision theory and utilitarian ethics. *The American Economic Review*, 68(2), 223–228.
- Kahneman, D. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, 47, 278.
- Kleiman-Weiner, M., Gerstenberg, T., Levine, S., & Tenenbaum, J. B. (2015). Inference of intention and permissibility in moral decision making. In *Cogsci*.
- Marr, D. (1982). Vision. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*.
- Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in cognitive sciences*, 11(4), 143–152.
- 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.
- Nichols, S. (2004). *Sentimental rules: On the natural foundations of moral judgment*. Oxford University Press.
- Parfit, D. (2011). *On what matters* (Vol. 1). Oxford University Press.
- Rawls, J. (2009). *A theory of justice: Revised edition*. Harvard university press.
- Rehg, W. (1994). *Insight and solidarity: The discourse ethics of Jürgen Habermas* (Vol. 1). Univ of California Press.

Scanlon, T. (1998). *What we owe to each other*. Harvard University Press.