Sociologica. V.14 N.2 (2020)
ISSN 1971-8853

Moral Philosophy or The Sociology of Morals? Response to Cansu Canca

Gil Eyal*

Submitted: August 23, 2020 – Revised version: August 24, 2020
Accepted: September 4, 2020 – Published: September 18, 2020

Abstract

I make three points in my response. I begin by pointing out the differences between the sociological and philosophical approaches to moral questions. The sociologist is interested in the trolley problem as a frame, and in the rhetorical power it generates. Second, I reject the claim that I am forcing the debate into a binary choice. Instead, I show the similarity between the model of moral reasoning Canca advocates and risk assessment, noting the well-known limitations of risk assessment. Finally, I reject the claim that I make moral arguments without engaging in principled moral reasoning, and instead explain the sociological method of comparison and relativization upon which I draw.

Keywords: Trolley problem; moral reasoning; frames; risk assessment; retroaction.

Acknowledgements

My thanks to Paul Appelbaum and David Stark, who provided some last minute advice and reminded me what really matters, from a moral point of view, in this debate.

* Department of Sociology, Columbia University (United States); ge2027@columbia.edu; https://orcid.org/0000-0001-7194-3864

1 The Trolley Problem

The function of trolley problems, Cansu Canca (2020) says, is not the one I attributed to them (Eyal, 2020). They were not constructed in order to present the solution to difficult moral problems as "a simple numbers game," a calculable trade-off, but as a device for "distinguishing the morally relevant underlying factors — such as intentions, causation of harm, duties, rights, moral claims, and the lives at stake — when making a moral decision." For example, you can ask whether it matters or not that by shunting the switch I do not *intend* to kill the lone person on the tracks, but to save the three people on the other tracks. The imaginary scenario can then be used to develop (or deconstruct) legal or religious defenses for problematic practices, such as the doctrine of "double effect" that justifies giving a dying patient morphine knowing full well that by doing so we are *also* hastening her death (Lavi, 2005). Canca is absolutely right to say, therefore, that the trolley problem can be used to support non-utilitarian or even non-consequentialist arguments.

Does it mean, therefore, that I "misapply and mischaracterize the trolley problem"? I think it is more the case that the sociologist's perspective here differs from that of the moral philosopher. These differences can be subtle because the two disciplines are, in fact, quite close to one another. They are interested in the same subject matter and ask related questions.[1] Both want to know, for example, what are "the morally relevant underlying factors," but a sociologist would never employ the trolley problem to answer this question because it is, as I said in my essay, a *frame*, a way of delimiting the relevant factors and imposing strong boundaries between what can and cannot be considered as relevant. The sociological style of inquiry is exactly the opposite. It typically proceeds by breaking down boundaries between things that supposedly could not be considered together (the Protestant ethic and capitalist economic action; cleanliness and morality; taxonomy and kinship; pasteurization and state authority; political philosophy and advice books). Sociology is driven to look at what was left "outside the frame," to multiply, rather than rarefy, the elements relevant to a problem. By the same token, the sociologist is less concerned with distinguishing between correct philosophical renditions of the trolley problem and laymen's versions. It is well-known that the trolley problem leads a very active and varied social life, far beyond the rarefied philosophical realm. The family resemblance between its different versions, I would argue, is better analyzed with the tools of the sociologist.

The trolley problem performs a particular function for the philosopher, Canca says. I was definitely a bit cavalier in not acknowledging how it can be used to "distinguish the morally relevant factors." But let's look more closely at how exactly it performs this function. It starts by setting up very strong boundaries as to what can be considered relevant. The original scenario is rigged so the consequences of one's action or inaction are fully determined (rails), fully knowable (the bird's-eye view), fully calculable (three is more than one) and nearly instantaneous. This means that really the only thing one needs to consider, the only element initially allowed inside the frame, are calculable consequences (and any choice other than full-blown utilitarianism would be moral cowardice, especially given the urgency that the framing creates). All the other beings that populate the realm of social morality are left outside the frame, congre-

1. This is clear from Steven Lukes' magisterial biography of Emile Durkheim (1973). Durkheim, the founder of French sociology, started out wanting to become an empirical moral philosopher. He envisioned an ambitious undertaking comparing the moral rules of different cultures. It was only in the course of pursuing this project, seeking to put moral philosophy on a firm empirical footing, that he discovered sociology. Durkheim's project to uncover the social bases of moral valuations was continued by the authors from which my essay drew inspiration — Mary Douglas and Luc Boltanski and Laurent Thevenot.

gating at the gate and clamoring to be considered. This is the starting point, equivalent to the mathematician's "let *x* be..." From then onwards, the philosopher plays the role of gatekeeper. She has time. She even has the time to modify how much time the other participants have. Bit by bit, she modifies the initial scenario and carefully inducts other elements, one by one, to discover the point at which they gain relevance (or not): at what point are the consequences no longer easily calculable? In comes the child and lays down on the tracks. At what point do intentions matter? In comes the fat man and climbs the bridge. What is the relevance of time? In comes Canca's magical lever that can slow down trains and switch back and forth between tracks. As gatekeeper, the moral philosopher is examining each element to see whether there is a way to include it while still approximating the clarity, parsimony and relative certainty of the original scenario. In comes the child, and now three is no longer more than one. But perhaps there is still a way to calculate and compare the consequences? We can "give them appropriate weights in a common unit of value, estimate how probable they are, compare them, and choose the policy that results in least harm." This metric can serve to defend the common moral intuition that the life of a child is worth more than the life of older people, and to integrate this intuition inside the frame in a way that still approximates the clarity of the original scenario.[2]

At this point, the gap between the moral philosopher and the sociologist rapidly widens. The sociologist asks *who* exactly is deciding how to assign weights to harms and benefits? *Who* creates the probability estimates? *Who* furnishes the information on which the calculation of estimates and weights is based? *Who* decides how much time we have? What if no good estimates or weights are to be had? And the sociologist also asks who is "them"? What do the ones thus weighted think about the weights and estimates assigned to them? How are they going to react when they find out? Canca is right that in my original essay I did not characterize the trolley problem in its full rigor. I am not a moral philosopher. As a sociologist, I was interested in the trolley problem as a *frame*, and in the rhetorical power that is generated by this frame, even when used by non-philosophers or when merely a whiff of it is added to the conversation. This means attending to what is left outside the frame (more about this in the next section), to what it does to time, and to the discursive position that the frame creates. The trolley problem creates a role that hovers above the fray, with a fairly clear line of vision to the consequences (so clear indeed, that any temporal lag between them shrinks and disappears); with the foresight that allows to estimate their probabilities with some confidence; with the means to weight and commensurate all the other actors and their claims, without being answerable to them. This position need not be occupied by the moral philosopher. Once created, others (psychologists, public health experts, commentators, politicians) can step in, take it over and exploit its rhetorical power for their own purposes.

2. Canca draws on Kamm (2020), who writes that it is wrong to defend this intuition by arguing that the elderly person has only a few more years left, and so she stands less to lose. The correct defense, she says, is to point out that the child is year-poor while the elderly person is years-rich. She says this is based on the principle that "priority should be given to the worse off." This is perplexing because the principle applies just as well if we are calculating who has more or less years left. If you have more years left, you are worse off by losing them. So why would Kamm insist so vehemently on the usefulness of the distinction between years left or years accrued? I think the value of the distinction for the philosopher lies elsewhere. It has to do with precisely the point I made above, namely looking for a way to approximate the clarity and certainty of the original trolley scenario. While there is some uncertainty about how many years the elderly person and the child will each live (we know this in the aggregate, but Kamm has just demolished what she calls arguments from "interpersonal aggregation," and recommended instead "pairwise comparisons"), there is no uncertainty about how many they've had up till now. Unlike years left, years accrued is a "common unit of value" that permits a definite, fully calculable comparison. This is an example of how the need to approximate the original scenario leads one to a position that looks very much like pure consequentialism.

2 Exit Strategies and Models

Canca thinks I misrepresented the graduated exit plans debated by public health experts and philosophers. They are far more nuanced and complex than a simple binary choice between "economy-first" or "lives-first". "Once these nuances are captured, the question does not fit into the structure of a trolley problem anymore." We agree on this point. In my original essay, I said that the public health experts and philosophers "are not advocating an abrupt end to the quarantine, but a graduated exit strategy conditional on improvements in certain key measures." Yet, Canca thinks that by failing to engage "the moral complexity of implementing an exit strategy," I am forcing the debate back into a binary choice between "economy-first" or "lives-first". Here I must disagree. After all, I wrote the essay precisely against framing the matter as a binary, trolley-like problem (and I *do engage* these moral complexities at the end of the essay, only Canca says that I do so "without moral justifications" — more about this in the next section). I doubt that readers of my essay would fail to notice that. I also doubt that anybody would deny that Republican politicians in the US have doubled down on a dichotomizing "lives or the economy" rhetoric. The main point in dispute therefore is limited to whether or not a dangerous whiff of this rhetoric has snuck also into more reasoned analyses.

I was certainly all too brief on this matter in my original essay. I can now take a little bit longer to explain what I meant when I wrote that "my sociologist's ears pricked up when I noticed that whenever Emanuel or Singer were challenged on one of the details, they resorted to some version or another of the tradeoff argument, offering a hypothetical calculation of number of life years lost." I detected a whiff of trolley framing in these responses, an effort to arraign on their side its rhetorical power, the sense of urgency it creates, and the way it positions the speaker. Comparing number of life years lost, for example, is typical trolley framing because the temporal lag between consequences (people dying of COVID-19 *now* versus "deaths of despair" in the *future*), precisely the time during which something presumably could be done about future consequences, is zeroed. Each time, they responded to an objection by sketching the outlines of a calculation in terms of some "common unit of value," weights and probabilities, that would justify opening up (gradually) as causing "least harm."

It is telling that Canca accuses me of a "refutation of risk assessment and moral reasoning." Indeed, in this type of trolley framing the two are equated. Moral justification becomes a form of actuarial modeling (Lakoff, 2015). Weights and probability estimates are calculated from past data, and where the data is lacking it is supplemented by assumptions and the trained judgment of the model-builder. Past data is used to transform uncertainty into manageable risk. There is a long discussion of the problems with this approach in my book about the crisis of expertise (Eyal, 2019, pp. 66–70), and I cannot repeat it all here. Suffice it to say that 1) past trends are not always good guides for what the future holds; 2) past data may contain systematic biases reflecting how it was collected; 3) the assumptions necessary to transform uncertainty into risk are influenced by various extraneous factors and are often disputable; 4) transforming uncertainty into risk means creating *ignorance* about all the factors not taken into account, the built-in biases of the data and the assumptions, etc.; 5) it also means creating genuine *indeterminacy* because whether the assumptions hold or not depends on the future behavior of relevant actors (Beck, 1992; Callon et al., 2009; Wynne, 1992).

None of this "refutes" risk assessment, but it should lead us to be cautious about when it is sensible to rely on it. I don't think every use of risk assessment is *hubris*. I *do* think that at this moment, in an unprecedented pandemic that has already surprised and foiled our collective efforts at prediction several times, it *is* hubris. In my book I note that there are cases when the

> assumptions and heuristics [necessary for risk assessment] are so well accepted and widely distributed that they stabilize the situation. They are shared widely by all relevant participants and are embedded in institutional routines and regulatory standards [...] Most importantly, it is possible to govern the actions of all the relevant parties [...] so they do not deviate significantly from what the assumptions anticipate. In these cases, it is probably proper to speak of risks rather than uncertainties (Eyal, 2019, p. 68).

I offered insurance as an example of such a case. It should be clear, however, that the current unprecedented pandemic is *not* such a case. The assumptions and heuristics of model builders (gradual exit strategies included) are not well-founded because the situation is unprecedented and dynamic. They are not shared and widely distributed because of political polarization, media cacophony and the crisis of expertise. They are not embedded in institutional routines and it is exceedingly difficult to govern the actions of all the relevant parties (unless you have already built an authoritarian state surveillance apparatus, or you are a relatively small and cohesive island nation). Fortunately, risk assessment is not our only option. The recognition that risk is often "a false friend" has led in recent decades to the development of alternative approaches to uncertainty, such as the precautionary approach (Callon et al., 2009), or specifically in the field of emerging infectious diseases, the approach of "vigilance" drawing on "sentinel" rather than actuarial mechanisms (Lakoff, 2015). Neither of these offer a foolproof blueprint for exit strategy, but both concur in erring on the side of caution, preparedness and humility about our ability to anticipate and predict the consequences of our actions — the procession of US universities currently discovering that their well-crafted plans for reopening have been undone by factors beyond their control, as well as by "animal house" parties, seems to bear them out.

This is at least one key lesson that I drew from the parable of Moses and the lamb, which Canca dismisses. If we are truly mindful of the limitations to our abilities to plan and forecast, then perhaps we should infuse our approach to formulating a gradual exit strategy with particular attention to the most vulnerable, those most likely to be harmed if our calculations prove faulty, and not with a futile attempt to calculate what would be the most efficient approach based on a dubitable "common unit of value" (more about this below). I did not suggest that every life should be saved at any cost, but rather that the mercy and kindness that characterized Moses's concern for the lamb should infuse our policymaking as well.

So this is why my ears pricked up (and, I admit, my blood pressure shot up too), when I read Emanuel's and Singer's reactions. When the uncertainties were pointed out, which should have given them pause, they replied by activating the trolley frame — the supposed urgency which justifies ignoring the uncertainties,[3] the zeroing of the time lag that allows to commensurate the consequences, the positionality of the model-builder who, presumably, controls what is included in the setup and when.

This latter point is important. Recent events in the US make it clear that virus models have a retroactive side, so to speak (Desrosiers, 2015; Esposito, 2013; Stark, 2013; Stark, 2020, pp. 74–77). Their predictions are incorporated into the calculus of the very actors that they are trying to model, who then act in ways that "retroact," act back upon the model and often undo its assumptions. This is why I said before that the sociologist is interested in *who* is "them," and

3. Canca participates in this framing. She twice says that "there is no guarantee of future [...] vaccine" in order to justify the urgency of implementing an exit strategy even in the face of profound uncertainties. Yet, if there is one thing that is fairly close to guaranteed in this unprecedented pandemic, it is that eventually there *will* be a vaccine, the question is only how quickly and how effective.

what *they* think about being estimated as "high risk" or "low risk," for example. The students who heard that they are low risk went out and partied, thereby becoming risky super-spreaders. This holds even more for gradual exit plans. Because they have profound redistributive consequences, allocating risk differentially, they face an imperative of legitimacy. They must gain the trust of those they model or risk being completely undone by indeterminacy.[4] This means that the model builder has to take into account also the communicative aspect of the model, what it says, how it is read, and by whom.

This is a second lesson I took from the rabbinical story about Moses and the lamb. Unlike the trolley problem, it is a model of moral leadership very much attuned to the communicative aspect. By trying to save even the littlest lamb, Moses (or rather the Rabbis and the Church Fathers who retell the story) signals to the flock a radical equality of distributive consequences. Not the "fairness" of the trolley problem, which communicates to the individuals that their fortunes depend on how they are "weighted," "estimated" and commensurated on somebody's ledger, but a transparent equality that is capable of mobilizing and aligning them. It effectively says: "the worthiest among us [Moses] will go to the utmost length to defend the least worthy [the little lamb], *ipso facto* we all have the duty to defend each other." Naïve or paradoxical it may seem from the point of view of balancing probable risks, this model is immensely savvier (witness its far longer pedigree and resilience) than risk assessment about the signal it sends, how it is likely to be read, and about the potential retroactive effects.

3 On Making Moral Arguments

Finally, Canca is perplexed about the arguments I offer at the end of my essay. "How," she says, "does Eyal reach any moral decisions *without engaging in any principled moral reasoning?*" (emphasis in the original [G.E.]). What she means by this is that my suggestions at the end — for example, that essential workers should be offered merit pay, guarantees of security of employment, and that existing regulations should be enforced to require their employers to provide them with as safe working conditions as possible — are offered "without any premises." Canca would like me to first state my premises — for example, the utilitarian's "greatest happiness of the greatest number" or Kamm's "priority should be given to the worst off" — and then rigorously derive from these my conclusions as to what should happen.

Here I must double down again on the difference between sociology and moral philosophy. Superficially, the moral philosopher and the sociologist ask the same question: what is the source of the authority of moral valuations? For Canca it is clear that the answer must be some clearly stated and well-defended premises, and a rigorous process of deduction from these. The sociologist does not share this method. As Mary Douglas says (1990, p. 111), the big moral decisions are made not by individuals, but by institutions. Accordingly, to answer the question about the source of authority of moral valuations one needs comparative institutional analysis.

What makes it so clear to us, for example, that the lives of children are worth more than the lives of adults or the elderly? The moral philosopher starts from this intuition, works her way back to discover the premise that "must" be underlying it — for example, that "priority must be given to the worst off" — and then works her way forward from the premise to construct a

4. This is why Bayesian approaches, which seem to offer a way out of the limitations of risk assessment, are also stymied. In Bayesian approaches, the priors (the assumptions) are provisional and are constantly revised in a reiterative process in light of new information. The problem then becomes one of trust. Constant revision of numerical targets that anchor consequential forms of redistribution leads fairly quickly to public mistrust and loss of legitimacy.

rigorous method of decision (whether counting number of life years accrued, or what have you). I put the "must" in quotation marks because the premise is selected on the basis of its usefulness for the methods of the moral philosopher, not because of some strict logical necessity.

Indeed, is the link back from the intuition to the premise so secure? Are we really considering who would be "worst off" when our guts clench and our hands tremble as the child is introduced to the trolley scenario? The utility or happiness or misery of the soon-to-be-dead individual — child or elderly person — is an altogether strange, incalculable quantity. Not even Schrodinger could write the formula for it. For the individual, this quantity is immeasurably large as long as she is alive, so large that quantitative distinctions are small change. Who can say whether lesser happiness is derived from another month with loved ones, another trip to a holy place, a last supper? Yet, this quantity is also absolutely nil when she is dead. Is it possible that what is really behind the intuition is something altogether different, namely *pain* — the pain of the people related to the dying individual? We judge this pain to be greatest for the parents of a child. We do so because we identify with them. The word "judge" is wrong. Identification means that our hands tremble and our eyes water and that even if we "judged," based on our "premises," that the right course of action is to pull the switch and let the trolley run over the child, we wouldn't be able to bring ourselves to do it because the intuition of pain is embodied. This intuition of pain is conditioned by institutions, namely by what we take to be the "right" order of things (children should bury their parents and not the other way around), and by the distribution of moral duties in society (the duty of adults is to defend children). So, when a child is introduced into the trolley scenario, our hands tremble not because the child's life is worth more, or because the child is the "worst off," but because we are conditioned to recognize the worthy individual as one who defends the lives of children and who takes care to protect those social links that are most emotionally intensive.

In a different society, with different institutions, moral intuitions might be quite different. When parents used to routinely bury their children, and the emphasis was on the moral duties of children towards their parents, and not the other way around, it would not have been obvious that the lives of children are worth more. Mary Douglas (1990, pp. 4–8) offers a variation on a classic, trolley-like, *topos* of moral philosophy: people stuck in a cave, the only way to survive until help arrives is to eat one member of the group. Is it justifiable for them to do this? How should they decide whom would be eaten? Should, for example, an older member be chosen because she has already accrued more years than the others? Douglas says that if the group is from a small, solidary village (what she calls a "hierarchical society"), young members of the group will be just as likely to offer to die in order to save the older and more important members. They would not be considering who will be the worst off, but the overall social order and how they fit within it: "never could they emerge into the light of day, having killed and eaten the squire or the parish priest or the scout leader." It turns out that behind our rationalist, individualist veneer, we are not all that different. We too are institutionally conditioned to consider how we fit within a particular vision of the overall social order when we overvalue the lives of children.

This all too brief exercise illustrates the sociologist's method when inquiring after the source of authority of moral valuations — comparison. The sociologist juxtaposes our moral intuitions to other intuitions held by other people. We all know that our predecessors did not necessarily share our moral intuitions, and often our contemporaries do not do so either. The intuitions are contrasted not in order to discover which is more valid, or what single premise can anchor all of them, despite their differences. The contrast is used as a critical lever, to relativize and defamiliarize what may seem obvious and taken for granted. This is what I did,

all too briefly, by contrasting the trolley frame with the Rabbinical story. I believe this is a valid method of "principled moral reasoning," but it is also the inverse of the method that Canca envisions. Instead of stating the premises and then deriving conclusions from them, this method proceeds by successively deconstructing a series of alternative, competing premises and each time recovering what each of them left outside the frame. The conclusions that can be reached in this way are not definitive deductions, but more like cautionary reminders. They are no less useful nonetheless.

References

Beck, U. (1992). *Risk Society: Towards a New Modernity*. London: Sage.

Callon, M., Lascoumes, P., & Barthe, Y. (2009). *Acting in an Uncertain World: an Essay on Technical Democracy*. Cambridge, MA: The MIT Press.

Canca, C. (2020). The Pandemic Doesn't Run on Trolley Tracks: A Comment on Eyal's Essay "Beware the Trolley Zealots". *Sociologica*, *14*(2). https://doi.org/10.6092/issn.1971-8853/11412

Desrosiers, A. (2015). Retroaction: How indicators feed back onto quantified actors. In R. Rottenberg, S. E. Merry, S. J. Park, & J. Mugler (Eds.) *The World of Indicators* (pp. 329–353). Cambridge: Cambridge University Press.

Douglas, M. (1990). *How Institutions Think*. Syracuse, NY: Syracuse University Press.

Esposito, E. (2013). Economic circularities and second-order observation: The reality of ratings. *Sociologica*, *7*(2), 1–20. https://doi.org/10.2383/74851

Eyal, G. (2019). *The Crisis of Expertise*. London: Polity.

Eyal, G. (2020). Beware of the Trolley Zealots. *Sociologica*, *14*(1), 21–30. https://doi.org/10.6092/issn.1971-8853/10842

Kamm, F. M. (2020). Moral reasoning in a pandemic. *Boston Review*, July 6. http://bostonreview.net/philosophy-religion/f-m-kamm-moral-reasoning-pandemic (Last accessed September 9, 2020).

Lakoff, A. (2015). Real time biopolitics: The actuary and the sentinel in global public health. *Economy and Society*, *44*(1), 40–59. https://doi.org/10.1080/03085147.2014.983833

Lavi, S. (2005). *The Modern Art of Dying*. Princeton, NJ: Princeton University Press.

Lukes, S. (1973). *Emile Durkheim: His Life and Work*. Stanford, CA: Stanford University Press.

Stark, D. (2013). Observing finance as a network of observations. *Sociologica*, *7*(2), 1–12. https://doi.org/10.2383/74854

Stark, D. (2020). Testing and being tested in a pandemic. *Sociologica*, *14*(1), 67–94. https://doi.org/10.6092/issn.1971-8853/10931

Wynne, B. (1992). Uncertainty and environmental learning: Reconceiving science and policy in the preventive paradigm. *Global Environmental Change*, *2*(2), 111–127. https://doi.org/10.1016/0959-3780(92)90017-2

Gil Eyal: Department of Sociology, Columbia University (United States)
https://orcid.org/0000-0001-7194-3864
ge2027@columbia.edu; https://sociology.columbia.edu/content/gil-eyal
Gil Eyal is Professor of Sociology at Columbia University. He is the author, most recently, of *The Crisis of Expertise* (Polity, 2019). This year, he is directing at Columbia University a Mellon Sawyer Seminar on *Trust and Mistrust in Experts and Science*.