Erkenntnis (2025) 90:2279–2296
https://doi.org/10.1007/s10670-024-00805-8

ORIGINAL RESEARCH

When Should Popular Views be Included in a Reflective Equilibrium?

Borgar Jølstad[1,2] · **Niklas Juth**[3] · **Carl Tollef Solberg**[1,2] · **Mathias Barra**[1]

Received: 29 March 2023 / Accepted: 29 February 2024 / Published online: 23 April 2024
© The Author(s) 2024

Abstract

It has become increasingly common to conduct research on popular views on ethical questions. In this paper, we discuss when and to what extent popular views should be included in a reflective equilibrium process, thereby influencing normative theory. We argue that popular views are suitable for inclusion in a reflective equilibrium if they approximate considered judgments and examine some factors that plausibly contribute to the consideredness of popular views. We conclude that deliberation and familiarity contribute to the consideredness of popular views, whereas prevalence does not. Furthermore, we raise the "fair trial" objection: that expert deliberation on popular views might fail to respect the consideredness of these views. Eliciting views approximating considered judgments directly from participants is therefore preferable to refining popular views or treating them as "bare" intuitions in the reflective process. These conclusions have important implications for which research methods will likely elicit popular views suitable for a reflective equilibrium process. Research focusing on the prevalence of views is unlikely to yield popular views suitable for a reflective equilibrium, whereas deliberative and in-depth work is more promising.

✉ Borgar Jølstad
borgarjolstad@gmail.com

[1] The Health Services Research Unit—HØKH, Akershus University Hospital HF, Sykehusveien 25 Postboks 1000, Lørenskog 1478, Norway

[2] Centre for Medical Ethics (CME), Institute of Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway

[3] Centre for Research Ethics and Bioethics, Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden

1 Introduction

One objective of applied ethics is to provide well-argued answers to moral challenges. Is euthanasia acceptable? How should we distribute scarce healthcare resources? Who are the *worst off*? Some challenges are particularly pressing because they involve practical dilemmas that must be resolved, and the solution must be justifiable to others.

It is increasingly popular to conduct studies eliciting popular views[1] on such moral questions. Recent examples include studies on trolley problems (Dahl & Oftedal, 2019; Greene et al., 2009), health inequality aversion (Robson et al., 2017), the choices of autonomous vehicles (Awad et al., 2018), and deservingness of access to healthcare during the COVID-19 pandemic (Chan et al., 2022; Reeskens et al., 2021). The collected popular views are typically referred to as *intuitions* (Greene, 2014) or *preferences* (Reeskens et al., 2021; Robson et al., 2017). A prudent question, then, is: Given information about individuals' moral views, how and to what extent should these views enter deliberation on normative questions?

One possible way of bridging the divide between the empirical and normative literature is via a reflective equilibrium process. Savulescu et al. (2021) have recently argued for a *collective reflective equilibrium process* to integrate popular views when assessing novel technologies. Baderin (2017) has also argued for a *public reflective equilibrium* where the public's contribution is their intuitions. Popular views seem particularly attractive and perhaps even morally relevant when we are searching for socially acceptable solutions to difficult ethical problems that affect whole societies.

The aim of this article is to discuss whether, and if so when, popular views are suitable for inclusion in a reflective equilibrium process.[2] We will argue that the suitability of popular views in a reflective equilibrium turns on the issue of whether popular views approximate *considered judgments.*[3] When we elicit more considered judgments directly from the population, we face views that we cannot easily ignore or discard in the reflective process. Considered judgments gathered from the population could then be entered into the reflective equilibrium process in the same way as judgments from the literature.[4] The resulting collective reflective equilibrium process will plausibly be better justified simply because it has considered, and incorporated or reasonably rejected, a broader range of considered judgments.

[1] We use the term *popular views* to refer to all views gathered from the public using any type of empirical research, be it quantitative, qualitative, experimental, etc.

[2] We aim to make our paper as uncontroversial as possible in terms of meta-ethical doctrine and therefore state our claims in terms of moral justification in line with the theory of reflective equilibrium rather than in terms of truth. What is required for our argument to be meaningful is that considered judgments matter for the justification of an ethical position; whether this is by leading us to truth, a more reasonable position, or something else can be decided by the reader in line with their theoretical leaning.

[3] We focus on the question of the consideredness of popular views. For a broader discussion of popular views and justification, focusing on healthcare priority setting, see Gustavsson and Lindblom (2023).

[4] Publicly elicited considered judgments should then have the same standing as other considered judgments. This also means that they, like other considered judgments, can be adjusted or discarded in the reflective process. Accepting that a public view is a considered judgment thus requires us to enter it into the reflective process but does not require us to accommodate it in the resulting reflective equilibrium.

This manuscript proceeds as follows: First, we discuss intuitions, reflective equilibrium, and considered judgments. We then (1) consider whether (some) popular views can justifiably be called *considered*, arguing that the concepts of *deliberation* and *familiarity* are relevant factors in this respect. Next, we (2) consider if the prevalence of an intuition contributes to whether it can be viewed as a considered judgment. The problem of (3) how to process 'morally untenable judgments' is then discussed before we (4) raise what we call the "fair trial" objection to either refining popular views into considered judgments or treating them as "bare" intuitions in the reflective equilibrium process. Finally, we (5) treat the issue of whether different methods of eliciting popular views are more likely to result in considered judgments in some detail. We conclude (1) that deliberation and familiarity contribute to the consideredness of popular views; (2) that prevalence, in general, does *not* make a view more considered; (3) that morally untenable judgments can be discarded because they conflict with deeply held moral convictions; (4) that the best way to handle the "fair trial" objection is to elicit considered judgments directly from the public; and (5) that elicitation methods that include opportunities for deliberation are more likely to provide considered judgments.

2 Preliminaries

A few preliminaries are in order before we proceed. First, we do *not* aim to settle the general question of whether intuitions are relevant to moral theory. If it turns out that we should base our moral theory on some form of foundationalism, then popular answers to moral questions are as irrelevant as other intuitions. This article is not arguing for or against coherentism (in the form of reflective equilibrium) or foundationalism as the superior theory of moral justification, but merely takes coherentism as the point of departure. We thus assume that reflective equilibrium is a valid theory of moral justification. Second, we do not take a stance on what precisely intuitions are. Our argument will be independent of any specific conceptualization of intuitions as long as the concept is applicable within the framework of the method of reflective equilibrium. What is important for our purposes is that popular views can be placed on a scale ranging from *"bare" intuitions* to *considered judgments*. Third, while we note that the prevalence of a view might be instrumentally significant to its democratic legitimacy, we will not consider whether such legitimacy is morally relevant in and of itself within a reflective equilibrium, all things considered. Fourth, we will not consider general questions about reflective equilibrium as a theory about the justification of ethical judgments (besides questions relating to the appropriateness of popular views as input to the process of justification); we instead refer to the extensive literature on these topics (See e.g., Singer, 2005; Strong, 2010; Sandberg & Juth, 2011; Tersman, 2018).

3 Intuitions and Moral Theory

Intuitions have, among other things, been conceptualized as beliefs, dispositions to believe, and *sui generis* states (Pust, 2019). We do not mean to take a substantive position on what an intuition is. All that we assume is that an intuition is a judgment that strikes us as *prima facie* plausible[5].

Intuitions in the above sense often play a prominent role[6]. For instance, a *case* is constructed around some dilemma or other choice situation and is presented for the readers' consideration. Intuitive resolutions are provided, and theory is formed or adjusted to accommodate these intuitions, or an explanation is provided for why the intuitions are not to be trusted after all (i.e., a debunking strategy is used). In another format, intuitive judgments are put forward as counterexamples to (what is initially considered to be) plausible theories. For example, the intuitive judgment that no number of transient, mild headaches should be averted instead of one avoidable fatality is presented as a difficulty for distributive theories that allow unrestricted aggregation of well-being, such as utilitarianism or prioritarianism (Carlson, 2000). Commonly, in the philosophical literature, the intuitions presented appear to be the intuitions of the authors. The authors, in turn, assume that these intuitions will be plausible to, if not even shared by, their readers. The intuitive judgments may therefore be the intuitions of a relatively narrow range of people: professional philosophers and others who read philosophical journals. The intuitions may, of course, be widely shared, but this often remains speculative.

This philosophical orthodoxy raises the following question: if we consider the intuitive judgments of philosophers as relevant to moral theory, are we then obligated to consider the intuitions of others similarly? If we consider intuitions important, perhaps we should consider whether intuitions are, in fact, widely shared, or at least not uncommonly held, in the general population. We should also perhaps search the population for intuitions that have not been considered in the literature.

This is an issue of moral epistemology, in the sense of viewing intuitions as evidence for or against moral theories or as an issue of being able to justify an ethical position to relevant others. John Rawls claims that justifying something to someone means "[…] to give him a proof of its principles from premises that we both accept, these principles having in turn consequences that match our considered judgments." (Rawls, 1999, p. 508). Knowing these mutually acceptable premises and considered judgments would then be central to justifying our ethical position.

[5] Cappelen (2012) has forcefully argued that intuitions are not as crucial for philosophy as many philosophers assume. While we believe that intuitions, at least in the form of prima facie plausible judgments, play a significant role in moral philosophy, it would likely be possible to formulate our thesis without reference to intuitions. We would then formulate our thesis using phrases such as "pre-theoretical judgments" or "what (spontaneously) matters to people."

[6] When discussing normative issues, as well as in other areas of philosophy. However, we believe that the status of intuitions on ethics is especially important because ethical theory generally attempts to do two things simultaneously: provide answers to what is morally right or good *and* make sense of our moral sensibilities. Rawls makes this explicit in *A Theory of Justice* (Rawls, 1999).

4 Reflective Equilibrium, Popular Intuitions, and Collective Reflective Equilibrium Processes

Reflective equilibrium is a theory of ethical justification that relies on considered judgments, coherence, and theory. This idea has implications for how we should justify moral judgments, i.e., moral theory (Sandberg & Juth, 2011). For instance, when working towards a reflective equilibrium, we are usually not interested in "bare" intuitions but instead seek support in *considered judgments*.

We are interested in considered judgments rather than "bare" intuitions for several reasons. First, "bare" intuitions, unlike considered judgments, are less likely to remain stable when exposed to counterargument or novel information. And, if someone would discard their intuition after some thinking, then that intuition seems less weighty as a source of normative content. Second, and in the same vein, considered judgments are more likely to represent the actual values of a person. We all have experience with coming to realize that our initial views cannot be defended after being exposed to new information or after a round of reflection and revising or rejecting initial emotional responses as unfounded. People ought to take our judgements more seriously when they are deliberate rather than whimsical. A third reason may be, depending on your meta-ethical commitments, that a reflective equilibrium inherits some of its justificatory power from the inherent credibility of the included propositions. Beginning with considered judgments rather than "bare" intuitions would be one way of increasing the credibility of the ensuing equilibrium. In all, the epistemic quality of a considered judgement compared to a bare intuition is likely to be higher, hence being a more valuable input in a process of reflective equilibrium. Focusing on considered judgments rather than "bare" intuitions also comes with the advantage of being in line with the Rawlsian orthodoxy on reflective equilibrium (Sandberg & Juth, 2011).

In the simple model of a *narrow reflective equilibrium*, considered judgments are weighed against ethical theories, and judgments and theories are adjusted until they cohere. The resulting coherent position is called a reflective equilibrium. In the more complex model, *wide reflective equilibrium*, considered judgments, ethical theories, and background theories are all weighed and made to cohere (Daniels, 1979). Background theories are not necessarily ethical theories and may include theories of human nature, rationality, cooperation, well-being, and (other) scientific or metaphysical theories. For instance, Parfit (1984) relies on a specific version of the psychological account of personal identity as a background theory when he argues that we have a duty to make things better for everyone rather than only ourselves. A wide reflective equilibrium is considered more justified than a narrow equilibrium, primarily because the former considers a richer body of evidence that may question considered judgments and ethical theories, making the outcome potentially more robust to objections. Importantly, as the example from Parfit shows, the theory and principles included in the reflective equilibrium should offer an explanation of our considered judgments; mere consistency is not satisfactory.

In their seminal accounts of reflective equilibrium, neither John Rawls nor Norman Daniels emphasized the process leading to a reflective equilibrium as a collective one. However, suppose a reflective equilibrium can justify an ethical position. In

that case, there are at least two straightforward reasons to conduct empirical research on moral intuitions and judgments in the population: we can examine whether some intuitions are (as philosophers like to assume) *widely shared*, and we can also unearth *novel* intuitions and perspectives not previously considered by philosophers. We will argue that the prevalence of an intuition, that is, to what extent it is widely shared, does not add to its moral weight. Examining the prevalence of an intuition will not provide us with morally interesting knowledge about that intuition's moral merit, although there might be independent reasons to empirically probe the epidemiology of ideas. An instrumental reason to care about popular views (regardless of moral epistemology) is that sampling the population can also point us toward what matters to people. This reason can also feed into the process of achieving a reflective equilibrium. The interesting intuitions will be those that philosophers have *not* already considered extensively, and such intuitions are likely to be intuitions grounded in specific experiences and expertise. As we will explore in Sect. 9, this will suggest whom we should try to elicit judgments from.

In Savulescu et al. (2021) collective reflective equilibrium process, popular views find their place in the reflective equilibrium as (filtered) intuitions or preferences. In addition to the classical problems with reflective equilibrium as a method of justification, the collective reflective equilibrium process imports its own idiosyncratic difficulty, namely, the status of popular views. Consider the *Moral Machine experiment*: Awad et al. (2018) sampled popular views on the ethical priorities of automatic vehicles by conducting experiments where participants from the general public solved whom-to-save dilemmas in traffic accident situations. Their results indicate that, in general, popular intuitions favor saving the young and the greater number. Awad et al. proceed to argue that their results should be considered when programming autonomous vehicles.

Harris (2020) disagrees and argues that aligning policy with mere preferences – without considering moral arguments, the law, rights, and duties – would be monstrous. Responding to Harris' objection, Savulescu et al. (2021) counter that neither uncritically adopting *nor* disdainfully disregarding popular views seems plausible. On the one hand, ethical theory should not disdain the public's views.[7] On the other hand, the fact that a view is widespread does not necessarily in itself make it ethically acceptable (more about this below). Following the collective reflective equilibrium process method, we can perhaps reach a moral reflective equilibrium between popular views and moral theory.

Savulescu, Gyngell, and Kahane argue that popular views are not (proper) considered judgments in Rawls' sense but that we can screen and filter intuitions to end up with reliable and robust data. On the assumption that intuitions are not a fortiori considered judgments in Rawls' sense[8], this has important implications for how we incorporate such views within a reflective equilibrium. If popular views are not con-

[7] It is important here to differentiate between experienced justifiability in the sense that the population considers something justified or legitimate and moral justifiability in the sense that something is morally justified. We are concerned with the role of popular views in moral justification. As already stated, we assume that the reflective equilibrium process confers moral justification.

[8] We use the term 'considered judgment' in line with Rawls' and Daniels' definitions, discussed below.

sidered judgments, we may defend a low threshold for disregarding them when they conflict with established ethical theory. However, if we can elicit properly considered judgments from the population, we would be bound to give these judgments more credence. Furthermore, distilling considered judgments from popular intuitions whenever possible seems beneficial, both because it is a way of respecting the population's views and because it would foster a more comprehensive, more robust, more considered, reflective equilibrium. We also believe this strategy would have several secondary benefits apart from achieving a wider reflective equilibrium. Incorporating popular views would facilitate justification of the reflective equilibrium to relevant others and could also mitigate democratic resistance.

5 What is a Considered Judgment?

> Considered judgments are simply those rendered under conditions favorable to the exercise of the sense of justice, and therefore in circumstances where the more common excuses and explanations for making a mistake do not obtain. The person making the judgment is presumed, then, to have the ability, the opportunity, and the desire to reach a correct decision […] the relevant judgments are those given under conditions favorable for deliberation and judgment in general.
> (Rawls, 1999, p. 42)
> A considered moral judgment, even in a particular case, is in many ways far more like a "theoretical" than an "observation" statement. […] Evidence comes from the way in which we support considered moral judgments as compared to observation reports: we readily give reasons for the moral judgments, and our appeal to theoretical considerations to support them is not mainly concerned with the conditions under which the judgments are made.
> (Daniels, 1979, p. 270)

These quotes from Rawls and Daniels illustrate a nuance in the concept of a considered judgment. Rawls' explication focuses on whether conditions are conducive to rendering good judgment. In general, we should put less credence on judgments where we are likely to be biased due to our actual position in society (e.g., prioritizing cardiac conditions before other maladies because of one's own hereditary risk) and more credence on judgments that are less susceptible to debunking. This approach is obviously related to the notion of an original position: we should have more credence in positions where we are likely to be impartial. Daniels focuses on considered judgments that are supported by reasons and theoretical considerations, and views conditions of judgment as of lesser importance. So, Rawls's account emphasizes that considered judgments have survived reflection and deliberation. On Daniels's account, considered judgments are the judgments that withstand theoretical scrutiny.

A considered judgment, as we see it, is a judgment that is the result of, or has withstood, deliberation and reflection with access to relevant information and counterarguments. This will necessarily also involve considering alternative positions;

in this sense, one could argue that the process of reflective equilibrium has already started. This may seem paradoxical in that the considered judgments are entered into the reflective equilibrium process. While the type of reasoning that goes on prior to and during the reflective equilibrium process is similar, reflecting on a judgment is likely to be more local than the process of reflective equilibrium. When, for example, reflecting on whether we should push a man onto the tracks in a trolley dilemma, we do not have to consider the full scope of theoretical problems facing utilitarianism and deontology, respectively. We most likely consider various similar scenarios to make up our minds. Still, the difference between the reasoning going on before and during the reflective equilibrium process is one of degrees rather than a qualitative difference.

6 When are Popular Intuitions Considered Judgments?

Returning to popular views and their place in normative ethics, it is interesting to consider to what extent and under which circumstances they would meet the standards for considered judgments. As mentioned in the introduction, popular intuitions and viewpoints are increasingly sampled using varying methods; surveys and questionnaires, qualitative research that engages more deeply with different people, citizen panels, and experimental economics are examples. It seems unlikely, on both accounts, that views elicited from the general population will be *as considered* as the views discussed in the philosophical literature, but this does not mean that they are not considered *to some extent*. The 'consideredness' plausibly comes in degrees, with whims and "bare" intuitions on one end of the spectrum and (more) fully considered judgments on the other. Recall that considered judgments have more credence than (mere) intuitions when constructing a reflective equilibrium. Therefore, it seems natural to conjecture that the closer a view is to being a considered judgment, the higher the threshold for discarding it.

Savulescu et al. (2021) argue that popular intuitions do not qualify as considered judgments in a robust sense, as we cannot assume that they approach Rawls' ideal judges. This position is plausibly related to the popular views considered in their article, which are mostly gathered using quantitative and experimental methods. If we could construe popular views as considered judgments, this would perhaps be the most straightforward way of integrating them into a collective reflective equilibrium.

6.1 Deliberation

It is common to hold that intuitions are ubiquitous in modern moral philosophy.[9] Intuitions are used as touchstones for ethical theory. Our intuitive responses to thought experiments and real-life cases are presented as "tests" of different moral theories. However, importantly, intuitions are seldom accepted at face value. Though intuitions are treated as possessing some prima facie credibility, they are often criticized,

[9] Some philosophers argue that we can do without intuitions in moral theory. While we cannot get into this subject in this paper, we tend to disagree. Se Sandberg and Juth (2011).

disregarded, and even contradicted. An intuition might be initially plausible, only to lose its appeal after considering its implications.[10]. Take the lives-over-headaches intuition already mentioned. The problematic nature of this intuition emerges after some theoretical work that shows the non-transitivity implied if we accept a limited aggregation view (Jølstad & Juth, 2022). Accepting the intuition forces us to consider fairly complex theoretical alternatives such as Alex Voorhoeve's (2014) Aggregate Competing Claims theory. Intuitions, while possessing a certain prima facie credibility, may require substantial work to be accommodated in theory and may be much more problematic than initially believed.

To be considered judgments, views must be subjected to *deliberation and reflection*. This qualification means that we should consider to what extent a view results from deliberation when deciding whether it qualifies as a considered judgment. If someone has spent a long time thinking about an issue, we could (and plausibly should) have a higher threshold for disregarding their conclusions than if a view is a mere gut reaction. Deliberating on our intuitions allows us to weed out intuitions that we, on reflection, are not particularly committed to and explore the limits of the intuition by considering new cases. Such deliberation should ideally involve considering the alternative (and conflicting) views, the ramifications of accepting the view in question (both in terms of theory and practical implications), and whether we are prone to bias regarding the issue. These ideal components of deliberation follow from two considerations we have already mentioned: first, intuitions are judgments that strike us as *prima facie* plausible. We do not know why these judgments strike us in this way; it could be because we are biased or because we have only considered or encountered a few relevant aspects of the issue. Deliberation, in this sense, involves looking closer and examining the robustness and soundness of our initial impression. Secondly, that considered judgments are somewhat closer to "theoretical" judgments than "observations" (see the above quote from Daniels, 1979, p. 270). In deliberating on an intuition, we delve into what it means and what it would involve to take it seriously. There is no other way to decide whether an intuition can transform into a considered judgment than to deliberate or reflect on it considering other judgments. Considering one's potential bias is particularly important; if we realize that we would not have the same view in another social position or if personal circumstances prevent impartiality, we should distrust that view. This is well aligned with how both Daniels and Rawls thought of considered judgments.

6.2 Familiarity

According to Greene (2014), our moral intuitions are less reliable in unfamiliar situations.[11] This unreliability is because "automatic settings can function well only when they have been shaped by *trial-and-error experience*." (Greene, 2014, p. 714).

[10] It is important in a coherentist theory of justification, such as reflective equilibrium, that no intuition is above revisability, contrary to foundationalism (Lemos, 2018).

[11] Greene uses "intuition" to mean an automatic response generated by "automatic settings". This is a more restrictive use of the term than commonly found in the literature on reflective equilibrium.

Greene's statement seems prima facie plausible[12]: personal experience matters for our reasoning. In a case of priority setting in health care, we are unlikely to be familiar with questions regarding large numbers of people, novel technologies, or the basis of moral standing. On the other hand, we are more likely to be familiar with situations involving primary health care, the importance of health for well-being, and the experience of caring for our loved ones. We believe an interesting parallel argument can be made regarding the relationship between familiarity and whether a view is more likely to approximate a considered judgment. As Greene argues, intuitions about familiar topics or situations have likely undergone a trial-and-error experience. That also often means having experience with the ramifications of the intuition in question. Familiar intuitions are also more likely to have been the topic of discussions and reflection in ordinary life, and consequently, people are more likely to be able to provide reasons for their judgments. Judgments about familiar situations are plausibly more considered because they are more likely to result from a deliberative process (as emphasized by Daniels) and because they have already survived such a process (as emphasized by Rawls). As this argument makes apparent, familiarity makes it more likely that a view approximates a considered judgment due to increasing the likelihood that it has been deliberated on; familiarity essentially functions as a proxy for deliberation. The various "most likely"-clauses in this paragraph do much work in this argument. Historically, many familiar intuitions would today be unacceptable. This is an issue we will deal with in Sect. 7.

A separate, but related, reason to believe that familiar intuitions are more likely to be considered is that familiarity with a situation allows us to appreciate more fully the nuances and concerns of a situation. This, in turn, makes is easier to reason about and deliberate on familiar phenomena.[13]

One potential difficulty with using familiarity as a proxy for consideredness is that there might be a conflict between familiarity and impartiality and other forms of bias. Recall that for Rawls, impartiality is essential when forming a considered judgment. Perhaps the situations that are familiar to us are also those where something is at stake for us? While there is perhaps a correlation between familiarity and partiality, we believe these are both distinct and generally possible to distinguish. Familiarity can thus function well as a proxy for consideredness so long as we also bear in mind that considered judgments should be as impartial as possible.

6.3 Prevalence

Does the prevalence of an intuition – that is, the number of people that share the intuition – matter for whether it approximates a considered judgment? An interesting possibility is that the prevalence of an intuition might say something about how many people have considered it. If an intuition is highly prevalent, it may be regarded as an indication of the fact that many have *retained* this intuition. Hence, the intuition has survived, and presumably, it has survived counterarguments and deliberations. Furthermore, any belief that any individual holds over time could plausibly be subjected

[12] Prima facie clearly often means the same as intuitive.

[13] We would like to thank an anonymous reviewer for suggesting this point.

to that individual's further and more thorough consideration. Hence, if an observed high prevalence of an intuition indicates that that intuition has been held over time by some individuals, this could mean that this intuition exists as considered judgment for some individuals. It would then not matter very much whether we sampled the intuition from someone who had herself deliberated the intuition. A diamond will cut glass whether the person we obtain it from recognizes its properties or not. If *somebody* has considered the view and preserved it against opposing views, surely that view would approximate a considered judgment.

However, there is the risk that (at least some) widespread moral intuitions are less considered precisely *because* they are widespread: it is easier to unreflectively maintain norms and values if they are shared by almost everyone else. They may be part of a paradigm in the word's original sense: a set of ideas so entrenched that we do not even consider them. If a part of the process that an idea must undergo to reach the status as a considered judgment is deliberation – including surviving counterarguments – then these entrenched intuitions would not qualify as considered judgments. Singer (2019), for instance, argues that common notions about the moral status of non-human animals are a moral blind spot, much like past days' racism and sexism, that can only be maintained since so many of us take them for granted.'[14]

Another hypothesis could be that prevalent intuitions, at least those formed freely, are likely to be familiar ones. Intuitions about unusual or new situations, such as trolley dilemmas or intuitions about novel medical technologies, are likely to be unfamiliar and less prevalent. If there is something to this line of reasoning, the prevalence of intuitions could attest to their consideredness. However, this reasoning is a form of affirming the consequent. While it follows logically from 'universal familiarity' – situations that all humans are familiar with – that intuitions surrounding these situations will be prevalent (and perhaps also considered), it does not follow that a prevalent intuition has become prevalent by this mechanism. Additionally, it is difficult to identify 'freely formed' intuitions. Many of our basic moral intuitions are likely to result from social pressures, regardless of whether it is the intuition that the Great leader is a moral role model in a dictatorship or the intuition that disagreements should be solved with discussions rather than violence in a stable democracy.

Hence, all things considered, we should probably reject the notion that prevalent moral ideas and intuitions are more considered just because they are prevalent. Nevertheless, prevalent ideas that we have reason to think also have been scrutinized by many are more promising candidates as being more considered judgments. However, they are so since they have been subject to deliberation. This may, for example, be the case when assessing the prevalence of intuitions among philosophers of a given field. If we believe that philosophers adjust their views according to the best arguments, the prevalence of views among philosophers may give us some insight into which views are more considered.

[14] Following up on the idea of a paradigm, it might be the case that views prevalent in earlier days can no longer be held unreflectively. If some hold these views today, they are more likely to approximate considered judgments because one cannot unreflectively endorse them. These would then perhaps be considered judgments that are nonetheless unacceptable (see next section).

7 Unacceptable Intuitions

An important issue remains: what if familiar or deliberated intuitions are simply unacceptable? Perhaps a subset of the population believes that we should prioritize healthcare resources based on racist, misogynistic, or other chauvinistic views? It is our considered judgment that some such views should be categorically discarded and that some standard of moral admissibility for intuitions to be accommodated in the (collective) reflective equilibrium process is necessary. One solution would be to regard all popular views as "bare" intuitions with low epistemic credibility. However, then they are, by definition, not considered judgments, and such a strategy would counter the entire enterprise of including *some* popular views that are considered into a collective reflective equilibrium. Alternatively, one might devise an admissibility criterion by classifying popular views as unacceptable when they conflict with other established considered judgments. A clearly racist but considered position could then be discarded because it is incompatible with very plausible ethical considerations, such as the equal moral standing of citizens. This strategy seems apt to screen out popular views that undermine important rationales for including popular views to begin with, like the equal standing of citizens or impartiality as to the source of an intuition.

This solution seems to run into a problem: it introduces theory too early in the reflective process. Such a concern aligns with Alice Baderin's (2017) argument for preferring the intuitions of non-philosophers in a reflective equilibrium due to their being unburdened by theory. Does not the idea of having intuitions as input in our reflective equilibrium become close to vacuous if they are compared to theory at the very outset of constructing the coherent system? We fail to see that this is a major problem on a coherentist epistemic account. Again, there must be a difference between "bare" intuitions (the first thought that comes to mind) and considered judgments. Moreover, how should one consider or reflect on a "bare" intuition (thereby making it a more considered judgment) unless one does it by reflecting on it in light of other judgments? This question is aligned with both Rawls' and Daniels' accounts of considered judgments. According to Rawls, one investigates if the judgment is flawed, say biased, which requires thinking about, for instance, whether one has a self-interest in accepting the moral judgment in question. That is, comparing the judgment with other propositions. Even more obvious with Daniels: the source of the judgment could be reflected upon, presumably in the light of other judgments.

This way of conceiving of a reflective equilibrium fits nicely with some judgments being discarded as unacceptable. Why are they considered unacceptable? Well, presumably because they contradict deeply held moral convictions, often of a general kind. Again, why do most of us immediately discard racist or sexist moral judgment? Because they contradict basic moral assumptions: they make a difference between persons where no morally relevant difference is to be found (our theory says). At most, this only shows that there are intuitions and considered judgments at different levels of abstraction that all enter the process of justification to begin with, which should be no surprise for proponents of reflective equilibrium (Sandberg & Juth, 2011).

Rejecting the deeply held moral convictions or assumptions would in general entail rejecting or substantially altering more wide general reflective equilibriums. While constructing a "local" equilibrium, such as an equilibrium on health care priority setting, we draw on previously accepted positions on broader topics such as justice or distributive fairness. While any judgment is in principle rejectable or alterable, there is a presumption to question the validity of parts of our moral outlook more than others. We can question the entire moral system simultaneously only on pain of moral nihilism and radical skepticism. The judgments we wish to reject as unacceptable are in contradiction with moral judgments, such as the judgment that we should treat like cases alike (i.e. the formal principle of justice), that are at the center of all the broader moral equilibriums. Whilethe formal principle of justice could, in theory, be rejected or revised according to the coherentist account of justification of moral judgements, the gains of such a costly maneuver must be great indeed. In other words, some parts of a theory are more at the core and need more justification than others, which aligns with the coherentist credo.

Along the same lines, we could argue that unacceptable intuitions are those that do not survive some test of reasonableness, as discussed by Rawls (1993;1999;2001) d Scanlon (2000). Rawls (2005, pp. xlii) argues that "Citizens are reasonable when, viewing one another as free and equal in a system of social cooperation over generations, they are prepared to offer one another fair terms of social cooperation (defined by principles and ideals), and they agree to act on those terms, even at the cost of their own interests in particular situations, provided that others also accept those terms." Demanding that considered judgments are reasonable in Rawls' sense would rule out sexist, chauvinist, racist, and otherwise bigoted views because these views do not respect the freedom and equality of others. From the perspective of Scanlon's contractualist theory, they would be unacceptable because relevant others could not reasonably accept them. A sexist proposal would, for example, be unacceptable to any reasonable woman. The tests of reasonableness have the advantage of providing a specific theoretical rationale and more specific guidance on which judgments are unacceptable. They do this by singling out a set of deeply held moral judgments and is thus compatible with our previous suggestion for weeding out unacceptable judgments.

8 The "Fair Trial" Objection

It seems possible to argue for refining popular views into considered judgments through expert deliberation after eliciting "bare" participant intuitions. Such expert deliberation would be an alternative to eliciting popular views that approximate considered judgments directly by gathering the views after proper deliberation and reflection. An advantage to this approach is that gathering "bare" intuitions is less resource-demanding than gathering views approximating considered judgments. However, requiring experts to adjudicate the consideredness of popular views raises a "fair trial" objection. In refining popular intuitions, theoreticians consider views that are not their own. Can we then be sure that we are giving the intuitions a "fair trial"? When constructing an argument, the philosopher ideally considers the strength

of various judgments and sometimes discards them due to theoretical considerations, e.g., consistency, or because other conflicting judgments outweigh them. We do not have the same access to the minds of the people whose judgments we are considering. A participant who holds the lives-over-headaches intuition might recant from this intuition when confronted with its implications – but they may not. In fact, they may have considered reasons for biting the bullet and be prepared to defend this stance. At the very least, we must acknowledge that the refinement process from "bare" intuition to considered judgment potentially discards popular views due to reasons that are not transparent to the people whose intuitions we are discussing, and to which reasoned objections may exist. Exposing participants to relevant objections during the deliberative process would be preferable to experts considering the same objections after gathering popular views.

The fair-trial problem also bears on the use of popular views as intuitions in a collective reflective equilibrium, as envisioned by Savulescu et al. (2021). Assigning popular views to this standing in the reflective equilibrium process again opens the possibility of not giving the intuitions a "fair trial" and discarding or refining views in ways the originator would disapprove of. This risk of disapproval is because we typically have a lower threshold for dismissing or refining intuitions than for dismissing or refining considered judgments. Eliciting more considered judgments directly from participants would entail a higher threshold for disregarding popular views, leading to a reflective equilibrium more aligned with the public's views. Eliciting views in this sense would involve ensuring that the views have been deliberated and reflected upon, for example, in a properly moderated group discussion where relevant objections are raised. This seems to align with the spirit of including popular views in the reflective equilibrium process. Therefore, we believe that eliciting more considered views directly from the population is preferable to transforming them into considered views via expert deliberation and assigning popular views the status of "bare" intuitions.[15] Eliciting these more considered views would involve acquiring views that have been the subject of, or are the result of, deliberation and reflection.

While the "fair trial" problem leads us to prefer to elicit considered views directly from participants – rather than refining intuitions – there will always be some interpretative work involved when working with other people's views.[16] The difference between refining intuitions and eliciting considered judgments directly is thus one of degrees when it comes to the risk of getting it wrong. This means that the philosophers (or other experts) constructing the reflective equilibrium should be mindful to a different extent than when working only with one's own views. Someone who sets out to construct a publicly informed reflective equilibrium takes it upon themselves to think *on behalf of* others. This incurs a profound responsibility that requires careful reflection. If one decides to work with less considered views, that are likely to

[15] This reason comes in addition to the reasons previously given for focusing on considered judgments (see page 5), it is not the primary reason for focusing on considered judgments.

[16] One anonymous reviewer argued that the "fair trial" problem can be avoided if the reflective process is genuinely collective, rather than a philosopher working with the views of others, and that this should be the ideal. We believe that philosophers, and other experts, are generally better positioned for the process of reflection, but concede that a collective reflective process would be the ideal in an ideal world where all relevant participants had the time, willingness, and knowledge relevant for the reflective process.

be worse approximations of what people would endorse upon due reflection, this responsibility is magnified due to the increased risk of getting it wrong.

9 Methods of Enquiry

There are various ways of eliciting popular views: experiments, surveys, qualitative interviews, mixed-methods studies, focus groups, citizen committees, and others (See, e.g., Bryan et al., 2002; Abelson et al., 2003; Greene et al., 2009; van Exel et al., 2015; Coast, 2017). These elicitation methods give us information with different strengths and weaknesses. One dimension along which the information elicited by these different methods vary is the dimension of *deliberation*. On the primitive end of this deliberation spectrum, we would find "bare" intuitions, unrefined by reflection and deliberation. On the other end, we would find judgments resulting from reflection and deliberation. The information from the "Moral Machine" experiment would arguably be at the unrefined end of the spectrum, whereas information from group discussions and citizen committees would be at the more refined end of the spectrum. The views at the refined end of the spectrum would be closer to considered judgments, making them excellent candidates for inclusion in a reflective equilibrium without much additional work. We can plausibly assume, at least if the group is suitably moderated, that people's views in these cases have been exposed to relevant objections and have been the subject of consideration. We should then have a higher threshold for discarding them in the reflective equilibrium process. Note that we can only make this assumption due to the possibility of deliberation that suitably moderated group discussions provide. We could also gather considered judgments from singular individuals provided we facilitate deliberation, but this would be a far more resource-demanding method of gathering considered judgments. The views on the unrefined end of the spectrum are plausibly some form of "bare" intuitions that would be modified if subjected to reflection and deliberation. Thus, these intuitions should be regarded with more suspicion, potentially disregarded because of their lack of coherence or justifiability. Views gathered with group discussion and citizen committees thus benefit from fitting in nicely in a reflective equilibrium.[17] On the other hand, the views on the unrefined end of the spectrum have the advantage of being numerous; data from the "Moral Machine" experiment includes the reactions of millions of people. Thus, they represent the general population to a degree that cannot be matched by the more labor-intensive methods such as group discussions and citizen committees. While we have argued that prevalence is not an indicator that a view is more considered, we might still have instrumental reasons for caring about

[17] Note that there are two questions here that remain unanswered: (1) to what extent focus groups and citizen panels can secure effective deliberation on different types of topics, and (2) if so, how to do this effectively. See Abelson et al. (2013) for an overview of research on public deliberation. Ideally, we should also know what kinds of biases are amplified or suppressed in focus groups. For example, Abelson et al. (2003) found that deliberation changed people's opinions, but more dominant views (highly ranked etc.) could become more entrenched with more deliberation. These findings seem perfectly in line with the process of considering various views; during deliberation, coherence is likely to increase in importance, with less important intuitions losing out to more important ones.

prevalence, and coupled with familiarity, such views are plausibly revisited. This does not speak to their acceptability, which must be assessed separately. A reflective equilibrium that considers the prevalence of views might, for instance, be seen as more democratically legitimate or more stable.

We have argued that prevalence is not a good indicator of the consideredness of a moral view. What remains to find are views that have not been treated in the theoretical literature. If the aim is to elicit such views, this has ramifications for whom we elicit judgments from. Novel judgments and perspectives are likely to be the result of expertise that philosophers lack or experiences that are not widely shared. This means that eliciting views in domains where we are confident that we have an overview of all the relevant views and perspectives is not a fruitful endeavor.[18] But we cannot know a priori whether we have such an overview, and the burden of proof must be on those that claim that popular views are unlikely to be of value in a given domain. Experts in emerging fields are also likely to have novel intuitions grounded in their knowledge and experience. Knowing where the useful information lies, of course, is difficult; it may be that the philosophical community is insulated from the judgments of non-philosophers in many fields.[19] In sum, the kind of enquiry that is likely to be useful for moral theory is qualitative work focused on gaining new judgments and perspectives. This is quite different from the kinds of studies that have gotten the most attention in the philosophical literature.[20]

10 Conclusion

The aim of this discussion was to better understand whether, and if so when, popular views are suitable for inclusion in a reflective equilibrium process. We have argued that deliberation and familiarity make it more plausible that popular views approximate considered judgments, whereas the same cannot be said about a high prevalence. Furthermore, we have argued that the "fair trial" objection means that we should be skeptical of the possibility of refining intuitions or treating popular views as "bare" intuitions in the reflective equilibrium process. Eliciting views approximating considered judgments directly from the public is, when available, the better option. These conclusions have important implications for which methods are appropriate when researching popular views and which topics we should focus on. Methods that allow for deliberation and reflection, such as focus groups and citizen panels, are more likely to result in views that are good approximations of considered judgments. This makes the results of such methods suitable for inclusion in a reflective equilibrium process. Methods focusing on the prevalence of views – such as the "Moral Machine" experiment – are not likely to result in views approximating

[18] Once again, there may be strong instrumental reasons for eliciting these views, such as democratic acceptability.

[19] See Rose et al. (2019) for an example where an assumed to be widespread intuition (on "pragmatic encroachment") is seemingly not shared at all outside philosophical circles.

[20] We would like to thank an anonymous reviewer for pressing us on the implications of our argument that we outline in this paragraph.

considered judgments and are, therefore, not as relevant as input to the reflective equilibrium process. However, there might still be instrumental reasons for caring about the prevalence of views, such as democratic legitimacy or the practical stability of the equilibrium.

Funding Open access funding provided by Akershus University Hospital (AHUS). The work of BJ and MB is supported by the Research Council of Norway, grant no. 303724. The work of CTS is supported by the Norwegian Health Directorate and partially also supported by the Research Council of Norway, grant no. 303724. The work of Niklas Juth is financed by Swedish Research Council grant 2021−01266. Open access funding provided by Akershus University Hospital (AHUS)

Declarations

Competing Interests Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abelson, J., Eyles, J., McLeod, C. B., Collins, P., McMullan, C., & Forest, P. G. (2003). Does deliberation make a difference? Results from a citizens panel study of health goals priority setting. *Health Policy*, *66*(1), 95–106. https://doi.org/10.1016/S0168-8510(03)00048-4.

Abelson, J., Blacksher, E. A., Li, K. K., Boesveld, S. E., & Goold, S. D. (2013). Public deliberation in Health Policy and Bioethics: Mapping an emerging, interdisciplinary field. *Journal of Deliberative Democracy*, *9*(1). https://doi.org/10.16997/jdd.157.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., et al. (2018). The Moral Machine experiment. *Nature*, *563*(7729), 59–64. https://doi.org/10.1038/s41586-018-0637-6.

Baderin, A. (2017). Reflective equilibrium: Individual or public? *Social Theory and Practice*, *43*(1), 1–28.

Bryan, S., Roberts, T., Heginbotham, C., & McCallum, A. (2002). QALY-maximisation and public preferences: Results from a general population survey. *Health Economics*, *11*(8), 679–693. https://doi.org/10.1002/hec.695.

Cappelen, H. (2012). *Philosophy without intuitions*. Oxford University Press.

Carlson, E. (2000). Aggregating Harms — should we kill to avoid headaches? *Theoria*, *66*(3), 246–255. https://doi.org/10.1111/j.1755-2567.2000.tb01167.x.

Chan, L., Schaich Borg, J., Conitzer, V., Wilkinson, D., Savulescu, J., Zohny, H., & Sinnott-Armstrong, W. (2022). Which features of patients are morally relevant in ventilator triage? A survey of the UK public. *BMC Medical Ethics*, *23*(1), 33. https://doi.org/10.1186/s12910-022-00773-0.

Coast, J. (2017). *Qualitative Methods for Health Economics*. Rowman & Littlefield.

Dahl, F. A., & Oftedal, G. (2019). Trolley dilemmas fail to predict ethical Judgment in a hypothetical vaccination context. *Journal of Empirical Research on Human Research Ethics*, *14*(1), 23–32. https://doi.org/10.1177/1556264618808175.

Daniels, N. (1979). Wide reflective equilibrium and Theory Acceptance in Ethics. *The Journal of Philosophy*, *76*(5), 256–282. https://doi.org/10.2307/2025881.

Greene, J. D. (2014). Beyond Point-and-shoot morality: Why cognitive (neuro)science matters for Ethics. *Ethics*, *124*(4), 695–726. https://doi.org/10.1086/675875.

Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, *111*(3), 364–371. https://doi.org/10.1016/j.cognition.2009.02.001.

Gustavsson, E., & Lindblom, L. (2023). Justification of principles for healthcare priority setting: The relevance and roles of empirical studies exploring public values. *Journal of Medical Ethics*. https://doi.org/10.1136/jme-2022-108702.

Harris, J. (2020). The immoral machine. *Cambridge Quarterly of Healthcare Ethics*, *29*(1), 71–79. https://doi.org/10.1017/S096318011900080X.

Jølstad, B., & Juth, N. (2022). Age and illness severity: A case of irrelevant utilities? *Utilitas, 34*(2), 209–224. https://doi.org/10.1017/S0953820822000024

Lemos, N. (2018). Foundationalism and coherentism in moral epistemology. In *The routledge handbook of moral epistemology*. pp. 375–386. Routledge.

Parfit, D. (1984). *Reasons and persons*. OUP Oxford.

Pust, J. (2019). Intuition. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2019.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2019/entries/intuition/. Accessed 2 August 2022.

Rawls, J. (1993). *Political liberalism*. Columbia University.

Rawls, J. (1999). *A theory of justice: Revised Edition*. The Belknap Press of Harvard University.

Rawls, J. (2005). *Political liberalism* (Expanded ed.). New York: Columbia University Press.

Rawls, J., & Rawls (2001). *Justice as Fairness: A restatement*. Harvard University Press.

Reeskens, T., Roosma, F., & Wanders, E. (2021). The perceived deservingness of COVID-19 healthcare in the Netherlands: A conjoint experiment on priority access to intensive care and vaccination. *BMC public health*, *21*, 1–8. https://doi.org/10.1186/s12889-021-10488-3

Robson, M., Asaria, M., Cookson, R., Tsuchiya, A., & Ali, S. (2017). Eliciting the level of Health Inequality Aversion in England. *Health Economics*, *26*(10), 1328–1334. https://doi.org/10.1002/hec.3430.

Rose, D., Machery, E., Stich, S., Alai, M., Angelucci, A., Berniūnas, R., et al. (2019). Nothing at Stake in Knowledge*. *Noûs*, *53*(1), 224–247. https://doi.org/10.1111/nous.12211.

Sandberg, J., & Juth, N. (2011). Ethics and intuitions: A reply to singer. *The Journal of Ethics*, *15*(3), 209–226. https://doi.org/10.1007/s10892-010-9088-5.

Savulescu, J., Gyngell, C., & Kahane, G. (2021). Collective reflective equilibrium in practice (CREP) and controversial novel technologies. *Bioethics*, *35*(7), 652–663. https://doi.org/10.1111/bioe.12869.

Scanlon, T. (2000). *What we owe to each other*. Harvard University Press.

Singer, P. (2005). Ethics and intuitions. *The Journal of Ethics*, *9*(3), 331–352. https://doi.org/10.1007/s10892-005-3508-y.

Singer, P. (2019). All animals are equal. *Ideals and ideologies* (11th ed.). Routledge.

Strong, C. (2010). Theoretical and practical problems with wide reflective equilibrium in bioethics. *Theoretical Medicine and Bioethics*, *31*(2), 123–140. https://doi.org/10.1007/s11017-010-9140-2.

Tersman, F. (2018). Recent work on reflective equilibrium and method in ethics. *Philosophy Compass*, *13*(6), e12493. https://doi.org/10.1111/phc3.12493.

van Exel, J., Baker, R., Mason, H., Donaldson, C., & Brouwer, W. (2015). Public views on principles for health care priority setting: Findings of a European cross-country study using Q methodology. *Social Science & Medicine*, *126*, 128–137. https://doi.org/10.1016/j.socscimed.2014.12.023.

Voorhoeve, A. (2014). How should we aggregate competing claims? *Ethics*, *125*(1), 64–87. https://doi.org/10.1086/677022.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.