

Better Minds, Better Morals

A Procedural Guide to Better Judgment

G. Owen Schaefer and Julian Savulescu

Abstract:

Making more moral decisions – an uncontroversial goal, if ever there was one. But how to go about it? In this article, we offer a practical guide on ways to promote good judgment in our personal and professional lives. We will do this not by outlining what the good life consists in or which values we should accept. Rather, we offer a theory of *procedural reliability*: a set of dimensions of thought that are generally conducive to good moral reasoning. At the end of the day, we all have to decide for ourselves what is good and bad, right and wrong. The best way to ensure we make the right choices is to ensure the procedures we're employing are sound and reliable.

We identify four broad categories of judgment to be targeted – cognitive, self-management, motivational and interpersonal. Specific factors within each category are further delineated, with a total of 14 factors to be discussed. For each, we will go through the reasons it generally leads to more morally reliable decision-making, how various thinkers have historically addressed the topic, and the insights of recent research that can offer new ways to promote good reasoning. The result is a wide-ranging survey that contains practical advice on how to make better choices.

Finally, we relate this to the project of transhumanism and prudential decision-making. We argue that transhumans will employ better moral procedures like these. We also argue that the same virtues will enable us to take better control of our own lives, enhancing our responsibility and enabling us to lead better lives from the prudential perspective.

Part I: Preliminaries

Introduction

Almost all of us want to be moral. Indeed, some maintain that simply entertaining claims about morality constitutively involves being (at least somewhat) motivated by those claims. And beyond being motivated, we also *care* – we want to do the right thing, to be a good person, to live a good and healthy life. The most difficult and contested of these concerns is the concern to be more moral, so we will address this "hard case". It is unclear how to go about being moral.

While some things are obvious – be kind, don’t steal, don’t lie, and so on – and automatic, other dilemmas require some thought and consideration. How much should one give to charity? What political party should one vote for? What products are too unethically produced to purchase? These dilemmas become even more acute in certain professional roles, where decisions affect institutional structure and large groups of people. Medical professionals deal with a wide variety of ethical issues, prompting many hospitals and research centers to form ethics committees. Consumers’ increasing ethical awareness has prompted interest in fair trade crafts and free range meats. And the justice system is built around judges and jurors deciding how to mete out fair and just punishment on criminal offenders. Ethical behavior then relies, at least in part, on moral reliability – that is, being able to reliably make and act on sound moral judgments.

The purpose of this article is to offer a guide on how to promote people’s moral reliability. Our preferred approach is procedural. This means assessing moral decision-making not in terms of outputs (what one decides or does), but rather the procedures involved in coming to that decision. We will delineate a set of 14 features central to decision-making.

This approach stems from an earlier article, Procedural Moral Reliability (Schaefer and Savulescu 2016). In that article, we derived six features of moral decision-making, inspired by the early work of John Rawls. These features were: logical competence; conceptual understanding empirical competence; empathetic understanding; openness to revision; and bias avoidance. In the present article, we will briefly summarize those features, and expand on them with eight more distinct capacities that would further improve reliability.

We take a procedural approach for two main reasons. One, the approach is more amenable to theorists with different normative commitments and should avoid many commensurate moral controversies (such as the classic disputes between consequentialism, deontology, virtue ethics and other normative theories). We do not presuppose which particular actions or beliefs are morally correct; instead, we look to personal capacities that, in virtue of their structure, are conducive to good moral decision-making. Two, we hold that there are indeed strong connections between certain procedural features of our thinking and the moral reliability of resultant beliefs and behaviors. These connections will be explored more deeply as we discuss each capacity in turn.

Notably, our approach focuses on the procedural capacities of agents themselves, rather than the decision procedures that make a given judgment reliable. Our present aim is not to establish a good procedure for determining the correct moral theory, or the answer to particular moral controversies. Rather, we are interested in the features of agents to be relatively morally reliable, which can in turn help answer the question of whether transhumans can be ‘designed’ to be more moral. And it will be somewhat easier to identify characteristics of agents that contribute to moral reliability than to lay out a correct decision procedure for agents to follow. Not only would such a decision-procedure inevitably invite the sort of controversy we hope this procedural approach would avoid, but it would require a much more systematic, step-by-step

account of how to work through a moral problem. With our approach, we can be more piecemeal. We do not have to offer a complete theory of the reasoning process, but can be content to suggest several core contributors to moral reliability that everyone can reasonably accept.

The Role of Good Judgment

But why should we care about good judgment anyway? Many ethical decisions in everyday life seem not to be the products of cool reflection, but immediate, intuitive responses. To be sure, well-reasoned decision-making is not the entirety of ethical behavior – however, we will argue that it is a significant part. This should come as no surprise, and indeed much literature on ethics in the past 3,000 years has inquired into just this issue. From ancient philosophers like Aristotle and Confucius to modern proponents of moral education and human enhancement, there has been enormous discussion of how to improve moral reliability. In this article, we will cull from insights both ancient and modern to offer a synthesized, broad-based picture of how to improve people's ability to make good, moral decisions to take us towards transhumanity.

To the extent that we care about doing the right thing and becoming more moral, we should also care about making good, reliable judgments. In fact, even immediate ethical decisions that involve no actual reflection or deliberation are the result of some sort of process – in such cases, an unconscious or non-reflective process. And oftentimes, we need to make split-second decisions. At the same time, for us to have any confidence in the morality of those quick judgments, we have to think that such processes reliably produce good results. And more importantly for present purposes, for any improvement to be possible, we have to be able to identify ways to make those procedures more reliable.

The features we delineate below, then, may not pop up in each and every moral decision we make. But they will be crucial in two ways: 1) assisting in explicit deliberation over moral issues; and 2) assist in explicit deliberation over how to ensure one can be most moral in cases where careful deliberation is not possible or desirable. In this way, good judgment can regulate moral decision-making in a wide variety of reflective and non-reflective contexts.

Part II: Cognitive Capacities

Logical, conceptual and empirical competence

Half of the Rawlsian features of moral reliability we previously identified were cognitive in nature. This is not a coincidence: insofar as correct moral judgments are internally consistent, conceptually coherent and empirically relevant, logic, conceptual understanding and empirical knowledge will facilitate those moral judgments.

Logical competence can assist in identifying internal inconsistencies, and exploring entailments of one's views. Inconsistencies are problematic insofar as they signal the need for revision of one's views – something has to give. On its own, logical analysis cannot tell you which view to abandon. Other capacities will be needed for that. But the first step of recognising a moral fault is necessary for it to be self-corrected. Entailments of one's views are similarly important, either so that one can live out more fully the substantial moral implications of one's considered moral views, or realise their absurd implications and be motivated to correct the views themselves.

Moral ideas are themselves complicated, and can primarily be grasped only at the conceptual level. This makes conceptual understanding important in moral decision-making; confusion or misapprehension naturally leads to errors of interpretation and implementation. We do not go so far as to claim this capacity involves having the *right* moral concepts (that would go too far afield from a procedural approach). Instead, we simply claim that one should have what Descartes called “clear and distinct ideas”. This can be coupled with logical competence, either to help identify errors and contradictions, or upon noticing an internal contradiction, be in a better position to grasp the weight competing moral considerations and thereby understand which should be abandoned.

But moral reasoning does not occur entirely in the armchair – one must engage with the world, and determine the best course of action within it. For this reason, being able to grasp the empirical reality of one's situation is also an important procedural feature. A number of ostensibly moral quandaries will in fact turn on such empirical claims. For instance, in order to generate the moral view that *we should combat climate change*, an empirical claim of the sort *climate change will cause significant harm* is necessary. Having the capacity to assess such empirical claims for accuracy will go a long way towards improving the reliability of one's judgments about practical matters.

Dealing with Uncertainty

Not all moral matters are black and white. Indeed, the sorts of cases where good judgment will be crucially determinative are just those cases where one is likely to lack certainty over what is the best thing to do. Even when one comes to a decision in the end, doubts may persist. Good judgment must take into account these doubts, and address uncertainty in a way that is both fair and productive.

The most straightforward aspect of dealing with uncertainty is accurate understanding of what one's credences (or degrees of confidence) actually are. The tendency to be overconfident in one's moral positions can have detrimental effects. Ignoring doubts can lead to discounting of evidence that might lead one to change one's view. Moreover, the strength of one's credences might well affect how one deals with moral issues. A high credence in a moral view might fairly imply one should strongly support a position, perhaps through strong, inflexible policies. If one

inaccurately reports certainty when doubts remain, this could lead to policies that are given too much strength and immutability. At the same time, reporting doubts when one's considerations really are decisive might engender a weaker policy than would otherwise be put forward.

But that might presuppose the answer to a prior question – what should we do when we don't know what to do? This question has received a significant amount of attention in recent philosophical work, and rightly so. Good judgment will require not just awareness of one's credences, but a reasoned plan of what to do when credences are at a certain level. Some decisions might only make sense when one is very confident – the changes are so drastic and risky that precaution against change is warranted. Other times, the opposite may be the case – even though one is extremely doubtful that some argument works, the relative downside of acting against the argument would be so massive that one should act in accordance with the dubious argument, just in case.

We will not here endorse one particular strategy for dealing with such uncertainty. However, we will suggest that good judgment will involve, in the very least, the agent addressing these issues thoughtfully and systematically applying a strategy for dealing with them.

Part III: Self-Management Capacities

Priority-setting

Moral judgments do not occur in a vacuum. One must deal with a variety of ethical issues in one's life, and not all issues can be addressed or examined. In addition to having reasoned consideration over the issues themselves, then, one will also have to address a higher-level question: how to manage the relative priority of competing moral issues and considerations.

At one level, this means forming an explicit understanding of one's values and priorities. What is more important, fairness or efficiency? Should one's family take precedence over the needs of the group? Are humans' interests strictly more important than those of other animals? Coming to grips with one's values in this way not only helps clarify what is at stake in a moral issue, but also allows one to come to a coherent justification for a given moral position.

But at another level, one must set priorities as to what moral questions to even consider in the first place. Much of this article is about explicit reasoning procedures which we can undertake and improve – but there are only so many hours in the day, and so much reasoning one can be expected to undertake. This is especially true in institutional contexts like meetings and debates, where issues must be decided within a set timeframe. So one aspect of good judgment is reliably picking out which questions to consider first and at greater length.

Procedurally, this means at least taking the time to decide what the most important issues are, reflecting on which are deserving of the most consideration and which must be passed over quickly. This might seem contradictory – more time spent on priority-setting will just offset any time saved later on. But if one's priority-setting is general enough, there will indeed be overall more time saved – a wide array of issues can be dealt with through a shorter categorical understanding of what needs attention.

Self-Control

Self-control (as we use the term) involves the tight connection between one's intentional actions and all-things-considered judgments. It is more or less the opposite of akrasia – akrasia involving intentionally acting against one's all-things-considered (ATC) judgments. Often, increasing self-control would be in someone's interest; say, a person recognizes that they need to lose weight but frequently succumbs to the temptation to eat sweets. In this chapter, it will be argued that akrasia reduction can also reliably make people more moral, at least in their behaviour and possibly motives.¹

Akrasia affects morality in a very straightforward way. Someone recognizes that some course of action is all-things-considered morally ideal or morally required,² but nevertheless fails to carry out that action. For instance, someone might recognize the moral imperative to donate significant sums of money to charity because that money could save a number of lives, yet remain selfishly tight-fisted. This is a failure of someone's consciously-held moral judgments to sufficiently motivate them to action. The person's ATC moral judgments are apparently overridden by an inclination that is not itself sensitive to all the relevant factors (self-interest may be morally relevant, but the ultimate disinclination to donate fails to take into account the overriding altruistic reasoning that is factored into the moral judgment that one should donate). By reducing akrasia, we could help ensure that those ATC moral judgments more often motivate action.

Straightforwardly, self-control contributes to moral reliability if ATC judgments are more reliable than their contrary inclinations. There are two motivations for that claim. Firstly, there is significant intuitive force behind the claim that ATC judgments can generally be relied on.³ When we attend to the general prospect of acting in accordance with one's ATC moral judgments and going against them, it is clear that the ATC judgments usually win out. If one

¹ One could even argue that similar means could be used to improve moral ideas. Rationalization, as discussed below, sometimes leads people's inclinations to affect and alter their ATC judgments; to the extent that this is a corruption of reliable judgments, reducing the effect of such inclinations on judgments could make those judgments more reliable.

² Cases of non-moral ATC judgments will be considered later on.

³ Jones (2003), who presses the point that one *sometimes* has most reason (including, sometimes, moral reason) to act akratically, concedes that usually, one does not have most reason to go against ATC judgments.

doubts one's ATC judgments, then it is unclear how one could coherently hold that one's view is correct, given all the evidence. Perhaps one could deny that one's moral views are ATC judgments at all, but this severely weakens any view. It means that one holds the view *without regard* to all relevant factors. One has left something relevant out, left it unconsidered when it could undermine one's view. And even if one has good reason to leave something out (perhaps due to time or resource constraints), it would be odd to claim that if everything had been factored in, the view would somehow become unjustified. To be sure, unlike in the case of basic intuitions, one is not committed to the reliability of *other people's* ATC judgments on these grounds – perhaps most people employ faulty reasoning that impairs their ATC judgments. But, in the very least people are committed to personal ATC judgment reliability, to the extent that the ATC judgment itself by definition involves a commitment to a particular judgment.⁴ Because of this, akrasia reduction would be a personal moral improvement.

A second reason to accept the general reliability of people's ATC moral judgments relies on the other sections in this article. ATC judgments, as a matter of course, involve the sorts of reasoning processes discussed previously that are conducive to coming to correct moral judgments. The 'things' in an ATC judgment are ultimately composed of basic intuitions, or more complex ideas (including other ATC judgments) built up from those intuitions. What's more, ATC judgments involve consideration of each of those factors, including how they hang together. That is to say, ATC judgments are the product of a reasoning process, which will display certain features conducive to reliable judgments. These include analogues to various facets of good reasoning: logical assessment, empirical assessment, attention to the content and strength of moral ideas in play, critical analysis of various factors, and (perhaps less commonly) identification of potential sources of bias. The presence of these processes alone does not ensure that ATC judgments are usually right. However, it does warrant the weaker claim (which is all we need here) that the ATC judgments are generally more reliable than the contrary inclinations.

Several sorts of interventions might help deal with akrasia. The most well-researched area of this is in treating addiction, which involves akrasia in the case of unwilling addicts who think, all things considered, it would be best to not to succumb to the addiction but do so anyway. For example, in the treatment of opioid addiction (including morphine and heroin), methadone is used as a relatively-benign substitute for more harmful substances; in combination with therapy, patients can be weaned off their drug habit. There is also some promise for using pharmacological interventions to promote impulse control. Impulse control is relevant to akrasia insofar as a common cause of akratic action are impulses that override better judgments.

⁴ It may be that one has higher-order reasons to doubt one's judgment. Perhaps one is aware of biases that corrupt it. In such a case, however, one is really forming a new ATC judgment – one that takes such biases into account and amends the judgment accordingly. One would then be committed to the reliability of the new (but not old) ATC judgment.

Bias Awareness

Cognitive biases distort reasoning generally, and as a matter of course they also distort moral reasoning particularly. As such, avoiding such biases is a key aspect of good moral reasoning.

But what do we mean by “bias”? This is surprisingly tricky, as the concept has not received a great deal of focused philosophical attention. Consider, for example, the understanding of bias proffered by Sunstein: “error in a predictable direction.” (Sunstein, 2005) This might match standardly understood cases of bias, such as racism: people will predictably favour their own race over others (which is a moral mistake). And it seems marginally more specific than error in general. But it is still far too broad. All predictable errors become biases, which includes errors that are predictable because they flow from some background theory. So, if utilitarianism is an incorrect theory, then utilitarians could be criticized as thoroughly biased in their views, in addition to being incorrect. This is not only an implausible application of the notion of bias, but seems like a form of double-counting of errors. Both the moral error and the bias appear to count against the utilitarian, when in fact the bias just reduces to that very error. In addition, it is a question-begging approach: it presupposes the correctness of certain moral judgments, and evaluates bias on the basis of deviance from those judgments. Any attempt to define bias in terms of erroneous results will run into this problem.

One could try define bias as a more narrow form of error so as to avoid improperly deploying it, but inevitably the concept will just collapse into a substantive claim over whatever the source of error is. Yet bias plays a more neutral role in debate. After being accused of bias, the interlocutor does not typically dispute that the bias in question is acceptable; rather, they deny possessing that bias at all. For example, it is not typical – on being accused of racism – to defend racism, but to deny that one is in fact being racist. This indicates that bias is a relatively neutral concept, or at least does not rely on substantively contentious issues. Yet if bias just reduces to a form of substantive error, we would expect many more such disputes than actually occur.

A more attractive understanding of bias along these more neutral lines comes from Nozick. Nozick, who thinks bias avoidance is an important component of rationality, suggests that first-order biases consist in the uneven application of standards. (Nozick 1993, p. 103) This is neutral insofar as Nozick does not define explicitly what those standards are. Bias, on this account, is really a form of internal inconsistency. Racism, for instance, is a bias because people on the one hand accept that race should not be taken into account in some context like employment but, on the other hand, those very same people do sometimes take race into account in hiring. This advantage may be somewhat subjectivist, but it captures why bias is so

universally accepted as problematic – bias occurs when people violate their own considered norms.⁵

Nozick’s definition, though, is still too broad. It implies that all forms of inconsistency are forms of bias, whereas the notion is generally used more narrowly. We should use a narrower definition of bias: *taking factors into account in a moral judgment that are not relevant to that moral judgment*. This is an intuitive account that does not include too much; utilitarians are not biased in applying their standards, as their standards are morally relevant. Even non-utilitarians should be able to accept this; one’s considered moral theory is surely relevant to how to act in particular cases.

We have used this definition to straightforwardly defend the moral importance of bias avoidance (Schaefer and Savulescu 2016). Briefly, ‘irrelevant’ factors introduce noise into the moral reasoning process. This noise tends towards arbitrary judgments, rather than well-reasoned ones that factor in considerations that are pertinent to the issue at hand. Bias avoidance consists in recognizing the source of this noise, and eliminating it where possible – or being so disposed so as not to be influenced by such irrelevant factors in the first place, as could be the case with transhumans.

Use of Decision Procedures

We have tried to suggest, at every point, strategies for dealing with particular flaws in one’s reasoning processes. But those solutions are typically tailored to the specific capacity in question. There is some reason to think, though, that a more systematic approach to certain problems can assist – especially with regards to multifaceted issues like bias. A growing literature has identified decision procedures that can mitigate some of these flaws.

Two features are needed for proper deployment of these procedures: awareness of what they are, and willingness to employ them. Awareness is a specific kind of empirical competence, one that pertains not just to the moral question at hand, but rather managing one’s judgments about the question at hand. We hope that this article can provide some insight into such management techniques, but it will of necessity be incomplete. Good reasoning will require further engagement with such decision procedures – reading up on what they are, and

⁵ Nozick, oddly enough, applies a more narrow and substantive definition to second-order biases, that is, biases in selecting the standards for first-order bias. Such second-order biases occur when “these very standards and weights would work to the exclusion or detriment of particular groups and this motivated them to put forward these particular standards.” (ibid., p. 103) It is not at all clear why Nozick only applies this motivational understanding to second-order biases and not first-order biases as well. Moreover, the second-order definition is too narrow – it excludes epistemic (rather than motivational) explanations of second-order bias, like salience, or any implicit second-order bias. It also abandons the advantages of neutrality, suggesting quite controversially that any standard motivated by exclusion or detriment of a group is biased. Correcting racism might be motivated by the detriment of whites, because in the racist mindset whites are unfairly privileged. On Nozick’s definition, then, correcting such racism would count as biased, which seems problematic. These factors make his second-order definition of bias unattractive, and so we will reject it.

experimenting with them in one's own life. Likely, any procedure will have to be tailored to one's own temperament and circumstances.

To translate that knowledge into practice, though, one needs to be willing to subject one's reasoning to a certain form of systematic management. This is somewhat controversial, but we will argue that it is both a reasonable attitude to have and incredibly helpful in actually bringing about reasoning improvements. We are, in many ways, deeply flawed reasoners – too close attachment to the way we are would foreclose any possibility of improvement. But that improvement will require a certain degree of management – recognition that the way we think about things needs to be altered somewhat in order to bring about better moral judgment.

Part IV: Social Capacities

Dialectics

Ideal arguments work to improve reasoning; they bring about moral improvement by exposing people to (among other things) factors they had not previously taken into account, overlooked flaws in one's own thoughts or ill-understood alternate ways of thinking. For this reason, an important component in promoting moral reliability will be dialectical skills.

Dialectical skills could roughly be divided into two categories: reception and communication. Reception involves the ability to fairly and accurately appreciate what others are saying in a discussion. Interlocutors will often have valuable insights that one had not previously considered, but misunderstandings will mitigate any potential benefit of the dialectic. Some of the other capacities, like conceptual understanding, might be of use here, but more generally one needs to be patient and understanding – as well as, at times, charitable to positions with which one disagrees.

Communication is the converse, fairly and accurately communicating one's ideas to others. Having one's ideas clearly formulated in one's head is an important first step, but practical ability to express those ideas effectively will go beyond mere clarity. Conciseness will be important – too long a discourse will turn people off and can lead others to miss your central point. Some linguistic expertise will be relevant – both to find the right words and phrases to encapsulate your ideas, as well as avoiding jargon that (depending on the interlocutor) may obfuscate. More generally, one should be able to read one's interlocutors to get a sense of whether they are following one's statements, and revise accordingly. And the medium itself – in-person speech, e-mail, online forums, etc. – will shape the norms of discourse, requiring one to alter one's communication methods appropriately.

When working in tandem, reception and communication naturally lend themselves towards improved reasoning processes. We often are overconfident in our own considered

judgments, and challenges can (if we are receptive to revision) help fill gaps in our own reasoning processes. Most exchanges will not result in one side or the other changing their opinion, but it can productively result in more rigorous reasoning processes and tweaks to positions based on pressing concerns and objections. And even if one is in broad agreement with an interlocutor, rich exchanges can help explore the reasoning process and illuminate the grounds for one's positions.

Empathetic Understanding

Empathy can be variably understood; here, as in "Procedural Moral Enhancement," we refer to the ability to understand and appreciate the state of mind of others. This allows one to more fully appreciate the personal interests at stake in various circumstances. In a way, this is a form of empirical competence, insofar as it is focused on internal phenomenology rather than the rules of morality or logic. Still, we can distinguish empirical and empathetic competencies by understanding the former as concerning facts about the external world, and the latter concerning people's inner experiences.

Straightforwardly, one can obtain empathetic knowledge by actually undergoing the experience in question. One is then able to extrapolate from one's own experience to that of others. But as noted by John Rawls, who advocated for the importance of empathetic understanding (equivalent to what he called sympathetic knowledge), "no man can know all interests directly." (Rawls 1951 p. 179) Moreover, for present purposes direct experience is not easily reconcilable with an agent-focused account of moral reliability. While we may be able to identify individual instances where a judge has had a given experience, it is difficult to identify a general reliable capacity of that sort. Perhaps being willing to undergo such experiences would make direct knowledge more likely, and a person's judgments more reliable in some cases – but given the impracticality of undergoing such experiences, it is likely such willingness would not contribute significantly to most moral deliberations.

Instead of direct experience, we can rely on mirroring or the ability to put ourselves in others' shoes. In this way, Rawls' referring to the capacity as sympathetic knowledge was inaccurate; sympathy typically implies certain emotional reactions like pity or grief. Those emotions may be important for other more substantive accounts of moral reliability, but they miss the core of the present account's reliance on the individual's level of psychological understanding. Empathy, which does not necessarily involve particular emotional reactions (though it is compatible with them), is more appropriate. Therefore, we refer to this competence as empathetic understanding.

The role of empathetic understanding is primarily interpersonal: when engaging in moral deliberation, we must typically consider the interests and feelings of others. Misapprehension of these factors may lead to ignoring potentially crucial considerations, or over-emphasizing the

importance of self-centred concerns. In this way, empathetic understanding can help avoid such self-serving biases. It is an important feature, insofar as the great majority of moral concerns will involve other agents, who have their own internal lives that may be difficult to scrutinise. A well-tuned sensitivity to others' concerns will go a long way towards ensuring responsible reasoning.

Imagination

In reasoning processes, one will not always have the benefit of earnest, well-articulated interlocutors to engage with or particular individuals' experiences to empathize with. This was an issue recognized by John Stuart Mill in *On Liberty* Chapter 2, and he astutely notes: "So essential is this discipline to a real understanding of moral and human subjects, that if opponents of all important truths do not exist, it is indispensable to imagine them, and supply them with the strongest arguments which the most skilful devil's advocate can conjure up." (Mill, 1999) Creative imagination of dialectics and individual situations will help ensure that good reasoning does not rely entirely on the good fortune of interaction with particular idealized individuals.

Dialectical imagination comes most directly from Mill, the process of imagining an interlocutor with whom to engage. If no one has brought objections to your view, come up with them yourself. Imagining someone diametrically opposed to one's view, and the arguments they would present, allows one to prepare adequate responses and revisions – even if no one on earth actually holds that diametrically opposed view. Charity will be even more essential here – one must be careful to avoid the temptation to construct straw men, and in earnest conjure up what is the most reasonable, convincing contrary positions. It may nevertheless be that the best counter one can imagine is thoroughly untenable, but that conclusion should only be drawn after a careful assessment of the fairly imagined arguments.

While dialectic imagination requires one to conjure up an interlocutor, empathetic imagination requires one to conjure up individuals who might be affected by one's moral opinions. When one is making a moral judgment, not all actual agents who may be affected are readily observable or knowable; some consequences may seem far-off, or uncertain. Some imagined reflection on the consequences of one's position will be needed. This may take the form of exploring potential side-effects of a policy in a practical manner. It may also require the formation of the tried-and-true philosophical method: the thought experiment.

Some may question the practical utility of thought experiments. They are sometimes criticised as obscuring the practicalities of an issue at hand: a fanciful thought experiment will isolate an entirely unrealistic set of circumstances, such that resolution of a problem in those circumstances will have little real-world meaning. Worse, one may worry that our judgments in fictional, sometimes absurd scenarios can have few implications for what we should do in more realistic settings.

Nevertheless, the thought experiment is a regular component of a philosopher's toolbox for good reason. They serve two essential purposes. One is to isolate extraneous influences and identify the core moral issue at play. Real-world scenarios contain a lot of 'noise' – multiple factors that all may be affecting the moral valence of a given scenario. Just as the most rigorous scientific experiments require controlling for extraneous factors, so must moral reasoning. By crafting examples that hold constant, say, the down-the-line consequences of an action, one can adequately address whether there is something intrinsically wrong (independent of down-the-line consequences) with some action.

The other purpose is to tease out all the relevant implications of one's argument. Some imagined, entirely fanciful and odd scenarios may nevertheless prove devastating for a position – such as Derek Parfit's Repugnant Conclusion as a challenge to total utilitarianism. (Parfit 1984) One may just bite the bullet in some of these scenarios, but one's moral judgments should at least be clear-eyed about its implications. And this feature feeds back into dialectical imagination: an imagined scenario can take the form of an argument against one's position that must be dealt with, one way or the other.

Part V: Motivational Capacities

Openness to Revision

Much of the preceding involves capacities that assist in prompting moral reflection or correction. But these capacities are useless if, upon discovering a fault, one is unwilling to revise one's views. Openness to revision consists in a dispositional attitude, where one is receptive to being wrong and needing correction. It is perhaps connected to the virtue of humility, though humility understood in the instrumental sense: humility may lead to willingness to concede error and change one's judgments (and, ultimately, behaviour).

The converse of humility is pride, and it is easy to see how pride can interfere with moral reasoning. Excessive self-confidence leads to discounting of others' views and ignoring of criticism, which may well be legitimate. This also has an inter-personal component, to the extent that being overly prideful makes others reluctant to engage, which will further shut off opportunities for moral improvement and correction.

But the more precise inverse of openness to revision, is stubbornness. Holding onto views, even when reason demands revision. Stubbornness can manifest itself in many aspects of life. It is particularly pernicious with morality, though, as there is not a natural 'check'. If one stubbornly refuses to believe friends' sound financial advice and make unwise investment, one may face the undeniable consequences of financial catastrophe, and be forced into correction (even if only after considerable loss). But moral mistakes do not have such a reliable external 'check'. Immoral action may lead to social consequences, but social views of morality are not

necessarily correct. Authorities may condemn, but then one may ask how one can be sure they possess moral authority.

This inscrutability of morality is part of why we adopted a procedural approach in the first place – absent a fully worked out moral theory we are confident in, we cannot properly verify that a given moral belief or action is finally correct. However, procedural features like openness to revision are in themselves verifiable – we can assess whether we are really open-minded in our views, or if stubbornness overtakes us when we debate and deliberate. And if there is a sensible connection between those procedural features and good moral judgment (as we have suggested), we can at least have some confidence in the output.

Recognizing faults in one's reasoning processes is not very useful if one does nothing about it. We also need to be open to revision. This can come in at a number of levels – being open to determine the implications of one's views, attend to reasons for and against, and change accordingly.

Willingness to Learn

Improving on various capacities like logic and empirical competence does not happen automatically. We require support from external sources, such as literature, newspapers and expert interlocutors. Sometimes, such as in schools, such capacities can be foisted on people. But among students, some will be more inclined to engage with the lessons offered and achieve improvement. And among adults outside the educational context, it is up to each person to self-motivate and seek out new sources of knowledge and assistance in deliberation.

Being willing and able to seek out such assistance is then an important part of achieving moral improvement. Like openness to revision, this works in tandem with the other capacities mentioned to improve reliability; given that those other capacities are indeed reliable, being motivated to take steps to expand on those capacities will itself be needed to achieve actual improvement.

In practice, this will mean different things in different contexts. For example, when deliberating over climate change policy, willingness to learn will encapsulate seeking out the relevant climate change experts and literature on the topic and being able to engage with it at the relevant level. When deciding how to spend one's money, it will involve reading up on how products are produced and listening to those who have voiced concerns. When evaluating who should get one's vote, it will mean attending to candidates' statements, as well as independent journalistic reports of their activities and commitments.

More generally, though, the thread that ties these together is a commitment to expand one's capacities when needed. Some of the recommendations in this article will require more

than a passing interest in moral improvement, and so one needs to have at least somewhat strong motivation.

Moral Courage

Pursuing the right course of action is often hard, especially when one's own interests are at stake. Holding an unpopular view may open one up to criticism, and require one to pursue an unpleasant or costly course of action. These downsides to a given considered moral judgment may tempt one to revise one's judgment, to make things easier. However, such considerations are, strictly speaking, irrelevant to the soundness of the judgment in the first place. If one really cares about coming to good, reliable moral judgments, one must have a certain degree of moral courage.

Moral courage, as we use it here, does not refer to the moral virtue of courage more generally – the sort of thing that encourages soldiers to sacrifice their lives on the battlefield. Such would be a substantive rather than procedural desideratum, and thus outside the scope of our proposals here. Rather, it refers to a more procedural aspect of our judgments – willingness to stick to our guns when under external pressure to change. While, as we have said, we should be open to revision based on competing evidence and moral arguments, the same does not apply to incentives that might be offered.

The question of demandingness might come up here. Just how much should we expect people to sacrifice in order to pursue good moral judgment? We cannot offer a specific answer, especially not within a procedural framework. It will be up to readers to decide how much they should care about good moral judgment over other things like social acceptance and prosperity. But we do recommend that people at least acknowledge the trade-off that must occur – that if they do bow to practical pressures, they not do so under the illusion that their moral judgment can remain sound and intact.

Conclusion: Prudence and Transhumanism

We have proposed a wide set of procedures that are, we argue, conducive to good moral judgments, and suggested ways to practically improve on them. These are meant to be applicable in a wide variety of contexts, both personal and institutional. With the framework, we believe that moral judgments will become more sound and reliable. This list we have provided is not exhaustive, but it is a start.

While we have maintained a proceduralist framework, it may be that in practice some substantive proposals end up being supported more by practitioners of the framework. This should not be surprising, and indeed it is to be somewhat expected if reliability is really increased. Also, it may be possible to weaken the proceduralism of our framework – to combine

our procedural approach with more substantive commitments, like valuations of altruism and justice, to achieve even greater levels of moral reliability. This is not incompatible with anything we have said here, and while such an approach will likely be more controversial, it is an avenue that is worth further investigation.

We have focused on the hardest case of how moral enhancement might be achieved. But the same qualities apply to prudential deliberation. After all, our future self is in many ways like another person. By applying these procedures in issues such as risk and health, for the sake of pursuing personal goals, one can arrive at better prudential decisions and take greater responsibility for one's own life.

Transhumanism is the project of humanity evolving to a higher level. One of the greatest achievements of humanity is the capacity to make normative decisions, both moral and prudential. But these capacities are also limited. In a globalised, technologically advanced world those moral limitations can render us unfit for the future (Persson and Savulescu 2012). We have provided a practical guide for humans can become fit for the future.

References

Jones, K. (2003) Emotion, Weakness of Will, and the Normative Conception of Agency. In *Philosophy and Emotions*. Cambridge: Cambridge University Press.

Mill, John Stuart.(1999) *On Liberty*. Edited by Edward Alexander. Peterborough, Ont.: Broadview Press.

Nozick, Robert.(1993) *The Nature of Rationality*. Princeton, N.J: Princeton University Press.

Parfit, D. (1984) *Reasons and Persons*. Oxford: Oxford University Press.

Persson, I., Savulescu, J., (2012) *Unfit for the Future: The Need for Moral Enhancement*. Uehiro Series in Practical Ethics. Oxford: Oxford University Press.

Rawls, John.(1951) "Outline of a Decision Procedure for Ethics." *The Philosophical Review* 60, no. 2 : 177–97.

Schaefer, G. O. and Savulescu, J. (2016) "Procedural Moral Enhancement," *Neuroethics*. Online First.

Sunstein, Cass R. (2005) "Moral Heuristics." *Behavioral and Brain Sciences* 28: 531–73.

Funding

Julian Savulescu was supported by the Wellcome Trust WT104848/Z/14/Z for work relevant to this paper