



Prudent Reflective Equilibrium



Denis Coitinho Silveira

(Universidade do Vale do Rio dos Sinos / UNISINOS, São Leopoldo; e-mail: deniscs@unisinis.br)
ORCID: 0000-0002-2592-5590

Abstract: The main aim of this paper is to propose the inclusion of the expertise of a prudent agent within the procedure of reflective equilibrium by adding a disposition for identifying reasonable beliefs. This can be seen as the starting point of the method, and would safeguard against the criticism of conservatism and subjectivism. In order to do this, I will begin by analyzing the core characteristics of the method and its main weaknesses. I will then investigate the characteristics of prudence as a disposition for identifying an adequate means for achieving a good end. With this in mind, I will apply prudence to the procedure which is carried out by an agent who deliberates well and can identify reasonable moral beliefs. These beliefs must be justified according to their consistency with ethical principles and with the factual beliefs of relevant scientific theories. Finally, I will argue that this deliberative process is consistent with ethical pluralism and democracy, and can be interpreted as a kind of moral knowledge.

Keywords: Reflective equilibrium; prudence; moral deliberation; moral knowledge; John Rawls.

Introduction

Reflective equilibrium (RE), as proposed by John Rawls in his *Theory of Justice* (Rawls 1971, 17–22, 46–53; Rawls 2001, 29–32), became the procedure par excellence both in normative ethics and in applied ethics, as well as in social and political philosophy¹. It has also been influential in the domain of law, especially in relation to Dworkin's theory of law as integrity. This is because it proposes that we leave aside controversial questions concerning the meaning of moral concepts and the truth of moral judgments, as well as about the existence of ethical properties, and identify moral objectivity in an inferential way. The basic premise of this method is that moral justification does not depend on an ultimate moral foundation, but on the coherence between all moral and non-moral beliefs,

¹ This article originally appeared in the Portuguese version titled "Equilíbrio Reflexivo e Prudência: um processo de deliberação moral", in *Trans/Form/Ação*, vol. 46(1), 2023, pp. 59–80.

namely, the moral beliefs, ethical principles and scientific judgments which are relevant to the issue at hand. The procedure is as the end point of a deliberative process in which we reflect on and review our beliefs. As stated by DePaul (2006, 618), “The best we can do is think things through and trust the conclusions we reach”².

RE, therefore, appears to encapsulate a non-absolutist normative model, since it is the end of a deliberative process during which reasons for and against are pondered and a particular course of action is chosen. This is done without reference to absolute moral standards such as the moral facts which would provide the basis of corresponding ethical principles, and which serve to justify a moral evaluation in a situation of uncertainty. It is a procedure in which we can test the correctness of a certain moral belief, or a group of moral beliefs, by means of our initial confidence in them, so that they become considered judgements. In turn, they may be coherent with one or more ethical principles, and perhaps with certain facts which have been confirmed by scientific theories, forming a wide RE (WRE). This is an efficient way of connecting values to facts whilst affirming that we can achieve objectivity in ethical decisions by means of internal consistence rather than truth. For example, we can justify the moral belief that “LGBTQIA+ individuals must have the right to form a family” based on its coherence with those ethical principles which consider gender discrimination to be wrong, and tolerance to be right. This is also coherent with the Constitutions in Western countries which affirm the equality of all members of society under the law and condemn all forms of discrimination, together with scientific descriptions of sexual orientation which have been undertaken by physicians and psychologists.

Despite the widespread influence of the above procedure, its effectiveness has often been called into question. The main criticism of RE has always centred on the fact that the result achieved depends exclusively on the considered judgements which constitute the starting point for the process. Philosophers such as Brandt, Hare, Lyons and Singer, for example, argued soon after the publication of *The Theory of Justice* (1971) that if there were no independent reasons for having trust in the starting point of the process, all that could be expected of the result would be internal coherence. At the heart of this argument was the initial credibility of moral beliefs, such as the fact that they were chosen only because people had confidence in them, and this confidence might reveal certain prejudices or biases which would imply conservatism. In addition, since individuals could arrive at a variety of coherent belief systems through RE, many philosophers considered that this procedure could be vulnerable to subjectivism (Brandt 1979; Hare 1973; Lyons

² Scanlon, Daniels and Walden are other influential proponents of the RE procedure, affirming that it is the only acceptable method for resolving moral questions, whilst other alternatives are mere illusion. All that remains is therefore the notion of normativity, since it is recognised that any absolutist model is completely impracticable (Scanlon 2003, 149; Walden 2013, 254; Daniels 1979, 256–257).

1975; Singer 1974)³.

In the light of the above, it is worth investigating how far the weakness identified in the procedure may be resolved by the inclusion of a virtue epistemology centred on practical reason or prudence. The main idea behind this is to rely on the disposition and ability of the prudent agent to identify the means required to achieve a good end, that is, the ability to deliberate well, which can mean arriving at reasonable moral beliefs. The advantage of this approach would be that our initial moral beliefs would become more reasonable, and the quality of our ethical reflection would become more sophisticated. This would provide an ability-willingness to identify reasonable beliefs, which would be taken as the starting point of the RE method, which could avoid the standard criticism about the lack of initial credibility of beliefs and the risk of relativism.

For this purpose, I will begin by analysing the central characteristics and main weaknesses of RE by focussing on the recent critique by Thomas and McGrath (2010), in which the authors argue that, even if the procedure is impeccably executed, it can lead the agent to hold unreasonable beliefs, thus leading to the conclusion that RE is an inadequate method. I will then consider the features of prudence as the disposition towards successful deliberation, in that it is an epistemic virtue which is connected to certain moral virtues. The next step will be to apply these features of prudence in RE so that the procedure may be executed by a proponent with sufficient practical wisdom to be able to identify a set of reasonable beliefs. Once these beliefs have been identified, the process of justification will take place by means of their coherence, both in terms of ethical principles and of factual beliefs. I think this would also represent an advantage for virtue ethics itself, which does not contain a procedure for justifying the decision of the prudent agent, a decision that is always taken to be the normative criterion of virtue. Finally, I will argue that prudent RE is consistent with both ethical pluralism and democracy, and may be considered as a kind of moral knowledge.

1. Explaining Reflective Equilibrium

Although RE is first presented in sections 4 and 9 of *A Theory of Justice* (1971), Rawls explains the method even more clearly in his article “The Independence of Moral Theory” (1975). In the latter, he defends the thesis of methodological inversion and states that progress in moral philosophy is independent of the study of the meaning of moral concepts and of the existence of moral properties, as well as the question of the truth of moral judgements and the question of personal identity. His general idea is to put the problem of objective moral truths temporarily to one side and analyse the substantive

³ A recent critique laid out by David Copp maintains that RE may imply a certain conservatism in moral theory, since the justification of beliefs would only become possible through the consistency between these beliefs and not through the truth of initial beliefs. The main problem lies in the determination of considered judgements which can only be identified by the degree of confidence of the proponent (Copp 1985, 141–149).

moral concepts which can be asserted in a specific situation. He therefore proposes to identify (i) a set of principles which are coherent with (ii) considered moral judgements and (iii) general convictions in RE. For instance, the principles of justice concerning equal liberty and fair equality of opportunity, in addition to the principle of difference, all of which were originally postulated under a veil of ignorance, would be coherent with our moral convictions that religious intolerance and racial discrimination are unjust, as well as with the ideas of cooperative society and personal reasonableness which can be found in several sociological and psychological theories. Thus, we can justify the principles in terms of their coherence with a coherent system of beliefs (Rawls 1975, 21).

The overall idea proposed is to consider specific moral judgements in comparison with general ethical principles and particular facts, then change them whenever incompatibility is observed in order to arrive at a consistent normative system. This process is indefinite, since new beliefs can always enter in a coherent system, thus leading to a revision of one or more of the existing beliefs. RE is therefore an example of a model of moral philosophy whose principal characteristic is its revisionist approach, in the sense that no moral belief is immune to revision, and any belief can be discarded if it is incompatible with new information or with a new system of coherent beliefs⁴. This evidences a marked distancing from dogmatism and absolutism by defending the idea of a moral agent as someone who is open to revising their beliefs and to seriously considering other points of view. It is important to point out here that this seems to be more suited to a project which seeks to identify a conception of justice for organising the basic structure in a democratic and pluralist society (Rawls 1971, 11–17).

Despite the great acceptance and wide use of this method in moral, social and political philosophy right up to the present day, a considerable number of critics have pointed out that the epistemological weakness of RE lies in the fact that it is not based on independent objective moral truths. One criticism which frequently recurs concerns the problem of the initial credibility of beliefs. It questions whether RE merely revises or reorganises positions which are already accepted in a society, something which could lead to conservatism. A further criticism of high importance is made in relation to the constructivist nature of the method, since it is considered that if the starting points of two or more individuals are different, they will then arrive at different points of equilibrium, so how is it possible to decide which equilibrium is more suitable? And does this not result in subjectivism, since there is nothing to rely on apart from the internal coherence of the beliefs of each agent?

If we look at these criticisms in more detail, we can first deal with the problem of the initial credibility of moral beliefs. Considered judgements are the starting point for RE, and these are beliefs in which we have great confidence and arrive at in an undistorted

⁴ As Rawls argues, "Reflective equilibrium requires only that the agent makes these revisions with conviction and confidence, and continues to affirm these principles when it comes to accepting their consequences in practice" (Rawls 1975, 8).

way. A good example might be a belief that “slavery is wrong,” that is, a belief based on intrinsic reasonableness, counting as a fixed point (Rawls 2001, 29–30). We then need to ask ourselves what makes moral beliefs reasonable, since it appears that simply giving credence to them is not sufficient to make them so. It is possible that a particular individual holds firmly to a belief but has arrived at it through a cognitive bias, namely, through tribalism, confirmation, or even availability. This may be a sign of ethical conservatism, in the sense that it is possible to identify one or more moral beliefs as “considered” only if they are part of the repertoire of the group to which a particular individual belongs. For example, someone may consider polygamy to be wrong if they have been brought up in a Christian religion, or if they live in the West, where the practice of polygamy is even a crime in several countries. As such, their confidence in the belief that polygamy is wrong may be simply a form of prejudice, just because the other members of their group consider it to be wrong and this confirms their own point of view. As Singer (1974, 516) points out, a considered judgement may only be a particular moral judgement which is made intuitively, and which may be based on rejected religious systems, prejudices, or on self-interest.

The second most common criticism of the method relates to the danger of subjectivism. The problem here is that there is no single RE shared by different individuals, since, depending on the range of considered convictions which are used as a starting point, we may arrive at different coherent systems of beliefs. This is the same as concluding that the method would not provide any moral knowledge, since knowledge appears to require moral truths. However, arriving at a set of beliefs (both moral and non-moral) and coherent moral principles would not imply truth, and thus one could be justified in believing a false belief or even a false theory. Singer (1974, 494), for example, identifies the danger of relativism in relation to morality and states that if one agent lives in a society and accepts a system of considered moral judgements, while another agent lives in a different society and defends a different system, very different moral theories could be ‘valid’ for both agents.

The problem becomes more dramatic when we consider that even individuals who live in the same society can arrive at very different belief systems. Let us imagine one person who is vegan (A) and another who eats meat regularly (B), where A thinks that “eating meat is wrong” and B is convinced that “eating meat is right.” Let us also imagine that A thinks we should respect all animals, both human and non-human, since all animals are sentient, that is, they feel both pain and pleasure, and therefore it would be wrong to be cruel to them. Further, we can imagine that A is concerned about the environmental problems which stem from the rearing of animals while, on the other hand, B thinks there is an important distinction between human and non-human animals, since humans have the capacity for reflection, and this shows them that cruelty to non-human animals is wrong but they can be reared and slaughtered in a “humane” way. Additionally, B may believe that animal protein is important for human health, and that there is no legal or

Christian prohibition of the rearing and consumption of meat (although eating pork, for example, is prohibited for Jews and Muslims).

We can therefore see that both the vegan and the meat-eater can arrive at coherent belief systems, even though their moral views are antagonistic. They are both justified in their positions because of the internal coherence of their belief systems, but this appears to be insufficient for us to know whether the consumption of meat is, in fact, right or wrong. In addition, we may wonder if this type of justification does not lead to a vicious circularity, even in the case of a wide coherent beliefs system, since a moral belief can only be justified by its consistency in relation to certain ethical principles and certain facts. But then the question is whether this form of justification does not result in arbitrariness⁵.

More recently, Kelly and McGrath (2010, 326–327) introduced a higher degree of specificity into the general criticism of epistemological weakness of the procedure by arguing that even if it is impeccably executed it may lead the agent to assert a set of beliefs which are not reasonable, and thus arrive at inadequate conclusions. The argument is presented as follows:

1. If RE were indeed the best method of justification, then the set of beliefs arrived at through impeccable execution could not be unreasonable;
2. Even if the method were impeccably executed, it would still be possible to arrive at an unreasonable set of beliefs;
3. Therefore, RE is not the optimal method.

The main point of the above criticism is that although an agent may apply the procedure impeccably, they could still arrive at a set of unreasonable beliefs because justification only occurs as a result of its consistency with a broader set of beliefs (Kelly & McGrath 2010, 346–354). It is important to point out that this criticism is a recurrent one in the debate on RE, and it is therefore not surprising that many philosophers maintain that RE should abandon coherentism and embrace a moderate foundationism which includes intuitionism, thus accepting considered judgements as true, since the result of this could then be interpreted as constituting moral knowledge (Ebertz 1993, 202–214).

I believe that one way of facing this problem and dealing with the criticisms levelled at RE would be to include the virtue of prudence in the method, so that a prudent agent could be able to identify reasonable beliefs – beliefs that are taken as the starting point of the procedure. The advantage of this proposal would be to interpret the outcome of this deliberative process as moral knowledge, which would be understood not as a true and justified belief but as an expression of certain virtues of the agents. In order to pursue

⁵ Siegel (1992, 43–44), for example, asserts that there is indeed a vicious circularity in RE, since the procedure does not provide any other element in addition to those which can be justified internally by their own coherence. He therefore maintains that the method fails as a plausible conception of justification for individual judgements as a result of these inferential principles.

this idea, however, we must first investigate the nature of the virtue of prudence and how virtue ethics and virtue epistemology might help us in this task.

2. The Virtue of Prudence

Prudence (practical wisdom) is an intellectual virtue, which signifies that it is a disposition to identify the most suitable means for achieving a good end. To put this more precisely, it is the ability-disposition to determine what will help an agent to reach this good end. It is related to deliberation, or it may be viewed as the capacity to deliberate well, since it is the ability to calculate and weigh up different reasons in order to decide on the optimal course of action to be taken. It is a virtue because it is a disposition of character or a trait of character, which is acquired through a process of habituation, that is, of constant exercise, and it is desirable since it is understood that the possession of it will guarantee, or at least contribute to, a successful life. It is intellectual or epistemic because, in the first place, it facilitates the identification of a good end, or at least demonstrates an ability to distinguish between a just and an unjust end and, more specifically, because it is a vital mental operation in evaluating the means by which a good end may be realised⁶.

It is important to recognise that prudence is only a virtue when it leads to a good end, since otherwise it would simply be an ability to calculate even when an end is wrong, and to understand that an end can only be virtuous when it is arrived at by appropriate means, since otherwise it would merely generate good feelings or a good intention. The procedure takes into consideration the consequences of an action and not ethical principles as absolute, and in addition takes into account common sense and what is common to humanity. For example, a prudent agent may decide that it is better to lie to someone who is unjustly persecuting an innocent person rather than tell them the truth. Clearly the agent knows when, in everyday situations, one should tell the truth, and that it is desirable to be honest, since personal prestige demands honesty and is essential to successful living. However, the agent also knows that to protect innocent persons is a question of duty, especially in cases of obvious injustice, as would have been the case with those who were persecuted by the Nazi regime during the Second World War. Thus, a prudent agent, after considering the reasons for and against the case, would then decide on the best course of action, that is, the one which is benevolent and just and is the correct decision in a particular circumstance. If telling the truth would lead to the death of an innocent person who is being unjustly persecuted, then it cannot be considered as a prudent one. It could be interpreted as an honest action but not as a virtue, since it would be against justice, benevolence and even courage. The prudent agent has the capacity to

⁶ Julia Annas (2011, 1–7) asserts that virtue involves a form of practical reasoning similar to that employed in carrying out a practical action, such as swimming, practising sports, or playing a musical instrument. As such, the virtuous agent's practical reasoning shares important characteristics with those of the practical expert, as long as this includes a specific aspiration on the part of the agent to become a better person.

correctly see what should be done in each specific case⁷.

It is important to remember that prudence was one of the four cardinal virtues both in Antiquity and in the Middle Ages. Aristotle, for example, when reflecting on prudence (*phronesis*) in Book VI of *Nicomachean Ethics* (NE), considered it to be a disposition which allows us to deliberate correctly about what is good and bad for human beings, not on goodness in itself but on goodness as it exists in the real world, and not in a general form but in certain specific situations. The agent then acts on the basis of this good deliberation. As Aristotle himself said: “Prudence, by contrast, is about human concerns, about things open to deliberation. For we say that deliberating well is the function of the prudent person more than anyone else” (Aristotle, NE VI, 1141b 10–12). In general terms, deliberation is the consideration of the reasons involved in resolving a difficult question which the agent needs to decide upon. The decision will be a good one if the desired end is achieved, that is, it will constitute a good action, and it will be bad if this is not the case.

Since we wish to rely on the expertise of a prudent person to identify reasonable beliefs, while considering the need to resolve the epistemological weakness of RE, I believe it is important to reflect on the essential nature of the virtue of prudence. Especially from an Aristotelian perspective, that is, one which is not Socratic or Platonic, prudence is regarded as a practical ability and not as a science. We should remember that, for Aristotle, ethics lies in the domain of practical rather than theoretical knowledge, and this means that it can only operate in accordance with proximate and outlined truths, since it deals with questions concerning what is just and good in situations of uncertainty and diversity (Aristotle, NE I, 3, 1094b, 15–25). Within this domain, we should consider prudence as the capacity to identify the relevant circumstances of each specific case. We are not seeking to identify true premises but are rather trying to choose between relevant values in order to decide on the best course of action, and so this approach may be regarded as a question of “knowing how” (practical knowledge) instead of “knowing that” (propositional knowledge). Looked at from this perspective, we must consider the experience of the agents concerned and the attention they pay to specific circumstances, since we can only deliberate about something when there is a choice to be made, that is, when no demonstration is either possible or sufficient.

It is also important to our central aim to understand how the different virtues involved are connected with a prudent action, since this is essential to our understanding of reasonable beliefs. Schematically, prudence is an intellectual virtue which is a condition for the possibility of moral virtues such as justice, benevolence and courage, because without this virtue it would be impossible to know what is really just, benevolent or courageous in a specific situation. Indeed, we would only know that it is simply desirable to be just, benevolent or courageous. Thomas Aquinas, for example, demonstrated that of

⁷ Hurthouse (2006, 285), like Annas (see note 5 above), understands that practical reasoning is an essential characteristic of the prudent agent, and refers to it as a ‘perceptual model’. What we may call a ‘perceptual model’ considers the special knowledge of *phronimos* to have the perceptual capacity to correctly see what one should do or how one should act in a given situation.

all the four cardinal virtues, prudence is the one which should guide the other three, that is, temperance, courage and justice. Without prudence, the other three virtues would not know what to do or how to do it, which implies that they would be blind or indeterminate. Without this deliberative capacity, just persons would love justice whilst not knowing how to achieve it, courageous persons would not know what to do with their courage, and temperate persons would not know how to reach temperance. In short, they would not know what constitutes the mean (*mesotes*). On the other hand, prudence without moral virtues would be empty, or at least it would no longer be an ability to calculate well, thus arriving at a vicious result, as would be the case nowadays if someone defended a totalitarian and dictatorial system as a way of achieving national prosperity. Without justice and benevolence, for instance, the agent could wrongly identify dictatorship as the most suitable means of achieving prosperity, whilst not considering how unjust it would be to deprive people of their freedom, and how maleficent it would be to censure, imprison or torture them. It is for this reason that the prudent individual must display a certain unity of virtues⁸.

Having made clear that this unity of virtues is fundamental to prudence, it is necessary to call attention to the deliberative calculus which can be understood as a practical syllogism where the major premise is represented by general or universal ethical principles, while the minor premise is identified in each specific case, thus leading to a conclusion which requires a corresponding action. Moral deliberation, which is the hallmark of the prudent agent, will then be concerned specifically with the minor premise in the practical syllogism. Let us consider, for example, that a prudent agent knows that courage is an appropriate means for his project to be successful (major premise), the next step would then be to identify if a specific act is one of courage (minor premise), and this would be followed by the knowledge that the specific act in question will be a way to achieve success, which will require courageous action. This refers us back to the case mentioned earlier in this paper, where someone who lies to an agent who is unjustly persecuting an innocent person is, in fact, committing a courageous act. It should be noted here that this is a deliberation which is sensitive to the context in which it occurs, since this practical reasoning implies that any adequate analysis of acts of courage requires considered judgements based on the experience of a mature agent. This requires a wide experience of the world and of us ourselves and of our obligations towards our fellow citizens, such that the decision deliberated must always take into account the consequences of our actions. The knowledge that lying to save the life of an innocent person who is being unjustly persecuted is indeed an act of courage is within the personal knowledge of the prudent agent, and this connects the virtue of prudence with the other

⁸ In his *Summa Theologica*, Aquinas asserts that the other intellectual virtues, such as wisdom, science and art, can exist without moral virtue, but this is not the case with prudence. This is because prudence is the direct reason for action, since it is preceded by judging and ordering in respect to its related means and ends, and this is not possible without removing the obstacles represented by the emotions (I-II, q. 58, a4).

virtues, such as courage, justice and benevolence⁹.

However, one problem with this kind of moral deliberation is that it lacks a justification process. Whatever the prudent agent regards as correct, that is, as the best course of action, is taken as a sufficient normative criterion for characterising it as a virtuous action. The limitation of this approach is that it is unverified and, as Aristotle himself pointed out, the prudent agent may err in their deliberations at some juncture. It therefore seems desirable to make use of a procedure to justify the moral beliefs and actions of the prudent agent. I believe that the RE method is appropriate for such purpose because of its fallibilistic nature, which appears to be very similar to the non-absolutist aspect of practical reasoning. If in addition to prudence and its connection with the relevant moral virtues we are able to rely on coherence as a normative criterion, I believe that this will constitute a methodological advantage, since we will be linking the knowledge of the prudent agent with a process of justification of beliefs which relies on the coherence of a decision with certain moral principles and with certain facts described by the specific sciences important in the case at hand.

3. Prudent Reflective Equilibrium

In the previous section we have looked at the characteristics of the virtue of prudence as a deliberative form of moral reasoning, but we have not considered to what extent a virtue ethics or virtue epistemology may assist us in perfecting the RE method, something which is imperative for being able to understand better the specific advantages I have in mind. I maintain that a virtue epistemology can be helpful because it explains knowledge in terms of true belief obtained as a result of the virtuous character of the agent, and not as a justified true belief. In this way, it will be to the credit of the agent, since cognitive success is attributed, at least partially, to their capacity for cognitive exercise. As one of the criticisms which is levelled at RE is that it cannot depend on truth but only on coherence for justification, if we think of knowledge as the exercising of specific virtues which allow the agent to have cognitive contact with reality, then this could be of great value in understanding better its epistemological complexity. I shall deal with this question in the final section of my paper by demonstrating that when RE is carried out by a prudent agent, it can be viewed as moral knowledge which is seen as a justified reasonable belief in WRE. My next task will therefore be to explain in detail how RE functions with the addition of the virtue of prudence, and I shall refer to this process as Prudent RE (PRE).

In a general way, PRE is centred on the expertise of an agent with the practical

⁹ David Carr (2020, 1391–1393) makes an interesting comment on this issue by pointing out that the knowledge of the prudent agent is closely related to the minor premise. He asserts that we should not deliberate about whether the agent should be forgiven when he/she possess appropriate excuses, since the agent has little difficulty in intuiting this fact. The specific case for deliberation would be to consider whether the guilty agent in a particular circumstance is, in fact, forgivable or not. This deliberative error would lead to the false conclusion that the agent should be punished.

wisdom to deliberate appropriately the means which are necessary to achieve a good end, and this can be viewed as arriving at a set of reasonable beliefs. It is important to mention here that this is vital in responding to the criticism of Kelly and McGrath (2010), who assert that even if the procedure is impeccably executed, it may lead the agent to propound unreasonable beliefs, which would therefore end in an inadequate conclusion. Once the prudent agent has established reasonable beliefs, the next step is to justify them through their consistency with the ethical principles set out by the main moral theories which are currently deemed acceptable in contemporary debate, and through their coherence with specific factual beliefs which have been affirmed by scientific theories accepted by their peers and relevant to the case being investigated.

But why should the good deliberation of the prudent count as having obtaining reasonable beliefs? Firstly, because the result which the agent achieves cannot be considered as a true belief, since the exactness of ethics is restricted to the practical circumstances involving diversity and uncertainty. The outcome arrived at is objective because it is associated with a science, albeit a practical one. Secondly, because the deliberation of the prudent agent, which is the choice of the best course of action that connects several virtues, can be interpreted as beliefs which reasonable agents would approve of, since reasonableness is a necessary condition for the possibility of the very harmonious co-existence between individuals. In other words, reasonableness represents a certain equilibrium between the agent's own reasons and those of others, without being excessive. In fact, reasonableness can be understood as a practical truth, and one which reasonable people would recognise as being adequate¹⁰.

If we consider, therefore, that the decision taken by a prudent agent cannot be deemed unreasonable, since the determination of the mean is its fundamental characteristic, the course of action chosen must be considered as equivalent to arriving at reasonable beliefs. The next step in PRE is to justify these beliefs through their coherence with a set of moral principles provided by the main ethical theories which seem to be accepted in the contemporary debate, such as utilitarianism, deontology, and virtue ethics. This step is important for checking if the decision taken by the prudent agent would be approved or refuted by these principles, or at least by the majority of them. The idea put forward here is that if a reasonable belief (or one which seems reasonable) is refuted by the principles of these moral theories, it must then be revised, since we know that the moral deliberation of the prudent agent can sometimes miss its target. The third step is for the agent to consider if these reasonable beliefs are coherent with certain factual beliefs which are supported by the relevant scientific theories for evaluating the

¹⁰ In a similar way, Scanlon (2014) defends the thesis that the considered judgements of RE should be understood as being clearly true. The point he is making here is that it is simply not sufficient to assert such judgements with confidence, since it is also necessary that the agent accepts that it is clearly true "when I am thinking about the matter under good conditions for arriving at judgements of the kind in question" (Scanlon 2014, 82). Thus, the justificatory force of judgements also depends on the substantive merits which can be identified during the process, and not only on their coherence.

particular case in question. This final methodological step is fundamental to guaranteeing that the ethical decisions made are linked with the real world, thus forming a more all-embracing system of beliefs. As with the second step, whenever there is inconsistency, the agent must review the initial belief. An example of PRE can be given by referring back to the dilemma described earlier concerning whether it would be right or wrong to lie in order to save the life of an innocent person who is being unjustly persecuted, considering a similar case to Nazism in the Second World War. This example helps to clarify the three methodological steps outlined above.

As we have already noted, the starting point for this procedure is the identification of reasonable beliefs by the prudent agent. Based on a good deliberation, the prudent agent concludes that it would be right to lie in order to save the life of an innocent person who is being unjustly persecuted. A specific example of this might be if a Jewish person was being persecuted by a Nazi agent during the Second World War, and the prudent agent knew where this person was hidden. On being questioned about the whereabouts of the Jewish person, the prudent agent's moral dilemma would then have been between telling the truth, or lying in order to save a human life. As this is clearly a situation of extreme injustice which is characterised by the arbitrary persecution and genocide of the Jewish people, the prudent agent forms an emotional connection with the person who is the victim of this injustice, and therefore, in an act of courage, lies to save his life. The conclusion to be drawn from this example is that even though there is normally a moral obligation to speak the truth, the situation in question demands that the prudent agent lie in the name of justice and benevolence. This is a reasonable belief because Nazism is intolerable for all the citizens who defend the equality of people and live in democratic societies.

We now need to justify this reasonable belief in terms of its coherence with a coherence system of beliefs consisting of ethical principles and scientific beliefs. If we begin with ethical principles, we can consider three moral principles based on the three main ethical theories which are accepted in contemporary debate, that is, the principle of welfare maximization (utilitarianism), the principle of universalizability and non-instrumentalization (deontologism), and the principle of the virtuous agent (virtue ethics), to see whether the belief that lying is permissible will be approved or not by the principles concerned.

According to the utilitarian model (act utilitarianism), the correct action is the one which maximises the welfare or happiness of those individuals involved. It clearly focuses on the best consequences of the act, namely, the best results, and does not regard any moral principle as absolute. In the case outlined above, telling the truth could lead to the death of an innocent person who is being persecuted by an unjust regime, whilst lying could possibly save them. If we take into consideration the potential results of telling the truth or lying, then the best action would obviously be the latter. It is, of course, possible to imagine undesired consequences of such an action; for example, the Nazi may discover

the lie and punish the prudent agent for it, which would raise questions as to whether it was indeed right to tell the truth. But considering the most probable results, lying would certainly be the best action to take.

In the deontological or Kantian model, the correct action is the one which would be approved by a rule which should be universalised, and which will not instrumentalize anyone, that is, a rule which does not just use people as a means towards an end. If we follow the formulation of the categorical imperative (first and second), it would be wrong to lie in any circumstance because we do not want the rule which approves of lying to become universalised. Thus, in accordance with the orthodox interpretation of Kantian ethics, the correct action would be to tell the truth whatever the consequences. However, if we adopt a less orthodox interpretation of the method, we may consider that it could be correct to lie, especially if we want the rule which approves of lying in order to save the life of someone who is being unjustly persecuted to be universal. In addition, we may call on the third formulation of the categorical imperative, if we maintain that lying is correct because it would treat the agent who is being unjustly persecuted as an end in itself and not only as a means. Yet despite the controversy over how the Kantian model is to be interpreted, we can use another deontological model which would clearly approve of lying, and that is contractualism. According to this ethical theory, as advocated by Scanlon, the correct action is the one which is approved by a principle which cannot be reasonably rejected. Thus, the principle which states that lying is wrong except in order to save the life of innocent people who are being unjustly persecuted would be justified, since it would not be possible to reasonably reject it, and this is because the principle which defends telling the truth even in the face of life-threatening persecution would not be acceptable for us¹¹.

Finally, in terms of virtue ethics the correct action is the one which is carried out by a virtuous agent who carefully deliberates on the appropriate means for achieving a good end. As we have already noted in the prudent agent's deliberation, the best decision would be to lie, for the sake of both benevolence and justice in the specific case referred to above¹².

In the light of all this, we can see that the reasonable belief, which is a considered judgement, that one should lie to save the life of an innocent person who is being unjustly persecuted, would be approved by the three moral theories. Or considering the orthodox Kantian model, it would be approved by most moral theories, and this fact could therefore

11 In Scanlon's contractualist theory, for example, reasonableness is determined by its acceptability to those involved. An action is considered wrong if it is prohibited by a principle which no one could reasonably reject, and this provides us with a direct connecting reason with other people's points of view. In this deontological model, the justification is intersubjective (Scanlon 1988, 189-247).

12 In "Virtue Theory and Abortion" (1997, 219), Hursthouse proposes a model of virtue ethics which is similar to the consequentialist and deontological models and is explained as follows: "P.1. An action is right if it is what a virtuous agent would do in the circumstances. P.1a. A virtuous agent is one who acts virtuously, that is, one who has and exercises the virtues. P.2. A virtue is a character trait a human being needs to flourish or live well."

be considered as a justification of the moral deliberation of the prudent agent.

Thus, the final step of the procedure is to see whether this reasonable belief would be coherent with specific factual judgements which are guaranteed by certain relevant scientific theories. For example, based on our knowledge of history, politics, anthropology, sociology and law, we can determine that dictatorial and genocidal governments are unjust. We can also give recognition to the value of democracy and of the rule of law as a means of ensuring the prosperity of nations. As a result, the reasonable belief under discussion would clearly be consistent with the factual beliefs referred to, and this could serve as a justification for its proximity to the real world.

We can therefore see that the method described above appears to be relevant because it allows us to use a procedure whereby we can justify the moral beliefs which we take as being reasonable. Let us imagine a situation where we consider that a prudent agent may defend a “reasonable” belief that “dictatorship is an appropriate method for achieving national prosperity.” If we apply the PRE method to this statement, this “reasonable” belief cannot be justified because it is clearly not coherent with the ethical principles of welfare maximization, universalizability and non-instrumentalization, nor even with the principle of the virtuous agent, and also because it is inconsistent with various scientific descriptions. History shows that totalitarian systems bring about national instability rather than prosperity, whilst the social sciences clearly demonstrate the value of democracy. Although it might be the case that a prudent agent could act outside the principles of the main ethical theories, and not be aware of the major historical facts and the most relevant contributions of the sciences, I believe this is extremely unlikely.

Conclusions. Reflective Equilibrium as Moral Knowledge

Now that we have clarified the methodology of PRE, which links the expertise of the prudent agent with a process of justification of beliefs through its internal coherence, we can consider how far the result of this process may be interpreted as moral knowledge. First, however, it is important to reflect on the problem of knowledge and explain what we understand as a virtue epistemology.

Traditionally, as in Plato’s *Theaetetus*, knowledge is understood as a justified true belief, and this means that for a person to have knowledge of a thing, this thing must be true, the person must believe it is true, and the belief must be justified, since it may be the case that someone has reached a true belief simply by luck. As an illustration of this, to have knowledge that there is a sheep in the field, the sheep must indeed be in the field, the person must believe that the sheep is in the field, and he must justify this belief through their own visual perception. The problem with this tripartite conception of knowledge is that, as Gettier (1963) points out, a person may indeed possess a well-justified belief which is true only as a result of luck. Let us consider a Gettier-type case. A person may profess a belief that there is a sheep in the field based on their visual perception. In fact,

this belief is true, because there is indeed a sheep in the field, but it is outside the person's visual spectrum, since the object seen is really a dog which looks like a sheep. Although the agent has a true belief that "there is a sheep in the field," which may be adequately justified by perceptual evidence, this is not a case of knowledge because the truth of the belief is obtained by factors which are aleatory to the agent's cognitive process.

Considering the problem with this traditional conception of knowledge, including the specific problem of moral knowledge, since concepts of right, goodness and justice, for example, would not exist in a world with natural objects, at least not in the same way as a sheep exist in a field, it would seem worthwhile to give consideration to the virtue epistemology. This is because knowledge is defined in terms of the agent's exercise of intellectual virtues. As Sosa (2011, 86) asserts, virtue epistemology is the conception that knowledge is the belief whose success is 'creditable' to the agent that exercise intellectual virtues.

The main thesis of virtue epistemology is that agents possess knowledge when they have a true belief which has been successfully formed through their own cognitive abilities or intellectual virtues, such as perception, memory and vision, from a reliabilist perspective – virtue reliabilism, or honesty, courage and humility, in a responsibilist model – virtue responsibilism (Roberts & Wood 2009, 6–9). The crucial point here is that when cognitive performance is unsuccessful, we do not attribute knowledge to the agent, and this demonstrates that the central focus is not on the analysis of beliefs but on the virtues and performances of the agents concerned. As Zagzebski (1996) correctly states, knowledge puts the agent in cognitive contact with reality, and does this in such a way that it can be characterised as good, desirable and even important. In the words of Zagzebski (1996, 270), "Knowledge is a state of cognitive contact with reality which arising out acts of intellectual virtue."

Bearing this in mind, PRE can certainly be viewed as a kind of moral knowledge, since it is the end of a deliberative process in which one weighs different reasons and then chooses the best course of action without the assistance of absolute moral standards, such as moral facts which would be the correspondentist foundation of ethical principles, and would serve as a justification for a moral judgement in a situation of uncertainty. Thus, their objectivity is ensured both by the prudence of the agents and by the internal coherence of their beliefs. They can rely on their capacity for good deliberation in order to specify reasonable beliefs, which in turn leads to greater precision in the method, since this process goes beyond their initial confidence and the adequate conditions by which they arrive at a considered judgement. This is because good deliberation is an action which weighs different reasons and chooses the best course of action, that is, the identification of the most appropriate means for reaching a good end. As prudence is a cognitive ability for achieving the desired aim which can be credited to the agent, it would be impossible to carry out PRE and arrive at beliefs which are not reasonable, since prudence is characterised by meeting the mean, which is still tested according to its

coherence with both ethical principles and scientific beliefs.

Thus, we may conclude that PRE appears to be equivalent to affirming a reasonable belief in WRE, and this may avoid the epistemological problems of conservatism and subjectivism, since the initial credibility of beliefs will be given by the epistemic capacity of the prudent agent to determine the mean and because the reasonable belief is tested not only by its coherence, but also by its proximity to the real world. In addition, PRE is highly consistent with ethical pluralism and democracy, being a kind of antidote to the dogmatism which is so threatening to diversity. This is because the procedure is always open to a revision of beliefs as a result of its internal coherence and its connection to the real world, and through the disposition of the agent to hit the target and to recognise the limits of reason. This deliberative model understands the importance of engaging in debate with other persons about various moral questions, whilst trying to understand ethical problems from a plural perspective.

But it is clear that the PRE has limitations. It would not be adequate for resolving complex cases in the private realm of morality, such as wanting to know whether it is right or wrong to “eat meat”, as in the question debated between A and B above. Here tolerance seems to demand that it is necessary to accept the various coherent systems postulated by agents, in terms of the different inputs which are considered legitimate within a democracy. However, in order to determine what is right and wrong in the public realm of morality, PRE does appear to be relevant to the identification of objectivity, since it includes in the deliberative model of the prudent agent a procedure for the justification of beliefs which is based on internal coherence, and this leads us to a clearly intersubjective point of view. The reason for this is that, in relation to the public sphere, it is possible to have the same public culture and common normative references as starting points (and these can be identified in constitutional matters and in judicial decisions, for example) for the assessment of specific problems. This would be the case in wanting to determine how the assets of a society should be fairly distributed, or how to justify punishment in an acceptable way, or even to specify the rights of minorities such as the LGBTQIA+ community.

Thus, having a justified reasonable belief in WRE, which so well expresses the prudent agent’s virtues within an interpersonal dimension, seems to indicate a more promising approach to the absolutist and dogmatic alternatives which consider truth, even in the moral domain, to be independent of human and social nature.

References

- Annas J. 2011. *Intelligent Virtue*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199228782.001.0001>

- Aquinas T. 1981. *Summa Theologica*. Trans. by the Fathers of the Dominican Province. Christian Classics.
- Aristotle. 1999. *Nicomachean Ethics*. 2nd Edition, trans. by T. Irwin. Indianapolis, Indiana: Hackett Publishing Company. <https://doi.org/10.1093/oseo/instance.00258595>
- Brandt R. 1979. *A Theory of the Good and the Right*. Oxford: Oxford University Press.
- Carr D. 2020. "Knowledge and Truth in Virtuous Deliberation," *Philosophia* 48:1381–1396. <https://doi.org/10.1007/s11406-020-00179-5>
- Copp D. 1985. "Considered Judgments and Justification: Conservatism in Moral Theory," in D. Copp & M. Zimmerman (Eds.), *Morality, Reason, and Truth* (pp. 141–169). Totowa, NJ: Rowman & Allenheld.
- 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>
- DePaul M. 2006. "Intuitions in Moral Inquiry," in D. Copp (Ed.), *The Oxford Handbook of Ethical Theory* (pp. 595–623). Oxford: Oxford University Press. <https://doi.org/10.1093/0195147790.003.0022>
- Ebertz R. 1993. "Is Reflective Equilibrium a Coherentist Model?" *Canadian Journal of Philosophy* 23(2):193–214. <https://doi.org/10.1080/00455091.1993.10717317>
- Gettier E. 1963. "Is Justified True Belief Knowledge?" *Analysis* 23(6):121–123. <https://doi.org/10.1093/analys/23.6.121>
- Hare R. 1973. "Rawls' Theory of Justice," *Philosophical Quarterly* 23:144–155. <https://doi.org/10.2307/2217486>
- Hursthouse R. 1997. "Virtue Theory and Abortion," in R. Crisp & M. Slote (Eds.), *Virtue Ethics* (pp. 217–238). Oxford: Oxford University Press. <https://doi.org/10.1515/9781474472845-016>
- Hursthouse R. 2006. "Practical Wisdom: A Mundane Account," *Proceedings of the Aristotelian Society* 106(3):283–307. <https://doi.org/10.1111/j.1467-9264.2006.00198.x>
- Kelly T. & McGrath S. 2010. "Is Reflective Equilibrium Enough?" *Philosophical Perspectives* 24(1):325–359. <https://doi.org/10.1111/j.1520-8583.2010.00195.x>
- Lyons D. 1975. "Nature and Soundness of the Contract and Coherence Arguments," in N. Daniels (Ed.), *Reading Rawls* (pp. 141–167). New York: Basic Books.
- Rawls J. 1971. *A Theory of Justice*. Original Edition. Cambridge, MA: Harvard University Press. <https://doi.org/10.4159/9780674042605>
- Rawls J. 1975. "The Independence of Moral Theory," *Proceedings and Addresses of the American Philosophical Association* 48:5–22. <https://doi.org/10.2307/3129858>
- Rawls J. 2001. *Justice as Fairness: A Restatement*. Ed. by E. Kelly. Cambridge, MA: Harvard University Press. <https://doi.org/10.2307/j.ctv31xf5v0>
- Roberts R. C. & Wood W. J. (Eds.). 2009. *Intellectual Virtues: An Essay in Regulative Epistemology*. New York: Oxford University Press.

- Scanlon T. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Scanlon T. 2003. "Rawls on Justification," in N. Freeman (Ed.), *The Cambridge Companion to Rawls* (pp. 139–167). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CCOL0521651670.004>
- Scanlon T. 2014. *Being Realistic about Reasons*. Cambridge, MA: Harvard University Press. <https://doi.org/10.1093/acprof:oso/9780199678488.001.0001>
- Siegel H. 1992. "Justification by Balance," *Philosophical and Phenomenological Research* LII(1):27–46. <https://doi.org/10.2307/2107742>
- Singer P. 1974. "Sidgwick and Reflective Equilibrium," *Monist* 58(3):490–517. <https://doi.org/10.5840/monist197458330>
- Sosa E. 2011. *Reflective Knowledge: Apt Belief and Reflective Knowledge*, Vol. II. Oxford: Oxford University Press.
- Zagzebski L. 1996. *Virtues of the Mind: An Inquiry into the Nature of Virtue and Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139174763>
- Walden K. 2013. "In Defense of Reflective Equilibrium," *Philosophical Studies* 166(2):243–256. <https://doi.org/10.1007/s11098-012-0025-2>