

# CONSCIENTIOUS SELF-REFLECTION TO THE RESCUE?

JOSHUE OROZCO & NATHAN L. KING

*Whitworth University*

Linda Zagzebski's *Epistemic Authority*<sup>1</sup> is an impressive, important, and wide-ranging book. We find much to admire within its pages. But in the spirit of philosophical interaction, our goal is to foster discussion of issues concerning which we find Zagzebski's treatment less than wholly compelling. We focus on (i) Zagzebski's assessment of the recent disagreement debates; (ii) the role of conscientious self-reflection in her solution to the epistemic problem of disagreement; and (iii) the broader role of conscientious self-reflection in her project. We argue that Zagzebski's notion of conscientious self-reflection is neither necessary nor sufficient for rational belief; nor does it provide the sort of cognitive guidance that is claimed for it. These considerations, we think, call for further clarification regarding the central role that Zagzebski gives to conscientious self-reflection. They thereby leave in doubt her specific solution to the problem of disagreement.

## THE PROBLEM OF DISAGREEMENT

Suppose you believe that  $p$ . You then come to learn that someone as reasonable and conscientious as you are, and who has the same relevant evidence you have, disagrees with you about  $p$ . How should you

---

<sup>1</sup> Linda Trinkaus Zagzebski, *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief* (New York: Oxford University Press). References to this book will be bracketed within the text.

respond? Recent discussion of this question has given rise to a spectrum of views. At one end of the spectrum are views on which learning of such disagreement is epistemically irrelevant. At the other end are views on which learning of such disagreement carries great epistemic significance – under such conditions, one should give equal weight to one's dissenters and abandon the belief that *p*.<sup>2</sup>

Zagzebski's diagnosis is that views on both ends of the spectrum stem from two competing and distinctively Modern values: the ideal of *self-reliance* (*egoism*) and intellectual *egalitarianism*. Views that accord no epistemic weight to a dissenter's opinion are supported by the ideal of self-reliance. According to such views,

It is *my* conscientiousness that counts for me, not hers. I trust *my* reasoning and other epistemic powers and not hers because my powers are mine and hers are hers. (p. 204)

In contrast, views that assign very significant weight to dissenting opinions are supported by egalitarianism:

When my belief conflicts with the belief of another person whom I judge has epistemic powers and virtues equal to mine, there is no reason to think that I am the one who is right, and so I have no reason to keep my belief. (p. 205)

Consider, then, the following positions regarding disagreement, both of which Zagzebski regards as 'extreme':

*Egoism*: In the face of peer disagreement, one should give one's dissenter's opinions no epistemic weight (and thus retain one's belief with unaltered confidence).

*Egalitarianism*: In the face of disagreement, one should give one's dissenter's opinions weight equal to that of one's own (thus moving from, say, belief to suspending judgment).

---

<sup>2</sup> For a seminal defence of the view that awareness of peer disagreement carries no independent epistemic weight, see Thomas Kelly, 'The Epistemic Significance of Disagreement', *Oxford Studies in Epistemology*, 1 (2005), 167–96. For seminal defences of the view that awareness of such disagreement is epistemically weighty, see Richard Feldman, 'Epistemological Puzzles About Disagreement', in Stephen Hetherington (ed.), *Epistemology Futures* (New York: Oxford University Press, 2006), pp. 216–36; David Christensen, 'Epistemology of Disagreement: The Good News', *Philosophical Review*, 116 (2007), 187–217; and Adam Elga 'Reflection and Disagreement', *Noûs*, 41 (2007), 478–502.

Zagzebski finds both options unpalatable. On the one hand, egoism seems to implausibly ignore – and treat as irrelevant – the fact that others are our epistemic equals. On the other hand, egalitarianism seems to lead to scepticism for many of our beliefs, and inappropriately ignores the unique role of the *self* in the resolution of epistemological problems. So what should we do? ‘The commonsense solution’, Zagzebski says, ‘is to compromise. Why not say that we should not place too much more trust in our own faculties than in those of others, but that we need not go so far as to think of others as our equals?’ Unfortunately, she thinks, this initially attractive solution is ‘actually hopeless’ (p. 206). She claims that while there is a principle argument for both of the extreme positions, ‘There is no argument for a compromise position on trust other than the desire to avoid the two extremes’ (p. 207). Indeed, she thinks, a middle-ground between the two extreme options is likely to be theoretically *worse-off* than the extremes themselves, because it may not itself enjoy support from either extreme.

Zagzebski is of course aware that middle-ground views exist. One such view is the *total evidence view*. On this view, the rational attitude to take with respect to some proposition in the face of disagreement is fixed by one’s total evidence, where this includes relevant *first-order evidence* (evidence directly relevant to the target proposition) and *higher-order evidence* (evidence about our evidence, or about our capacities for or performance in responding rationally to our evidence).<sup>3</sup> Thus, on total evidence views, the attitude that is rational to hold in the face of disagreement will depend both on the evidence that is directly relevant to the disputed proposition *and* evidence concerning the disagreement itself.

Crucially, advocates of the total evidence view attempt to give principled reasons for a compromise between egoistic and egalitarian values. Note that the character of the relevant first-order evidence makes no appearance in either of the extreme views. According to egoism, all that matters is my conscientious judgment; and this judgment matters because it’s mine. On egalitarianism, all that matters are the opinions of epistemic agents – namely, myself and my dissenter. But according to the total evidence view, it is a mistake to neglect first-order evidence

---

<sup>3</sup> For a defence of the total evidence view, see Thomas Kelly, ‘Disagreement and Higher-Order Evidence’, in Richard Feldman and Ted A. Warfield (eds.), *Disagreement* (New York: OUP, 2010), pp. 111–74.

in these ways. Indeed, in the absence of disagreement, such evidence is *paradigmatically* what makes our beliefs rational. It would be surprising if its relevance simply vanished once disagreement reared its head. By focusing on such evidence, in addition to higher-order evidence, we can see how total evidence views offer a principled middle way between egoism and egalitarianism. For perhaps in all cases, the higher-order evidence that someone one trusts believes  $\sim p$  is epistemically relevant to some extent. But perhaps in some cases, the relevant first-order evidence supporting  $p$  is so strong that it makes belief that  $p$  more rational than denying or withholding – even when the higher-order evidence is accounted for.

It is initially unclear why Zagzebski is dismissive of middle-ground views, including the total evidence view – especially in light of her own solution to the problem of disagreement. That solution features prominently the notion of conscientious self-reflection: ‘Given the argument of this book, it is reasonable to resolve the conflict in favor of what I trust the most when I am thinking in the way I trust the most, that is, conscientiously.’ (p. 214) When thinking conscientiously, Zagzebski says, an agent considers both *theoretical reasons* (e.g., arguments), and *deliberative reasons*. Deliberative reasons are irreducibly first-personal reasons that connect *me* to the truth of  $p$  (p. 64). Such reasons include items like intuitions and experiences. Crucially, they can also include instances of trust in others (p. 65). But notice: these latter items are *higher-order evidence*: they are evidence about whom I should trust to have good evidence, and to evaluate evidence rationally. Thus, Zagzebski’s view entails that the conscientious thinker will consult both first and higher-order evidence in the face of disagreement – a position that seems quite close to the total evidence view.<sup>4</sup>

The above remarks can be seen to yield a trilemma for Zagzebski: (i) accept egoism and inherit whatever problems attend it; (ii) accept egalitarianism and inherit whatever problems attend it; or (iii) accept that there is a principled middle way between egoism and egalitarianism – contrary to her initial judgment. Inasmuch as Zagzebski herself seems

---

<sup>4</sup> In this way her view is in agreement with total-evidence views. However, given the central role that she gives to conscientious self-reflection it seems to us that the character of the evidence, independent of what the conscientious agent takes it be, drops out of the picture. In this way, then, her view is much more closely aligned with an egoist position. More on this below.

reluctant to adopt (i) or (ii) – she offers lengthy considerations against both egoism and egalitarianism – we suggest that her best alternative is to adopt (iii). We provide additional support for this suggestion below. This is no objection to Zagzebski's solution to the problem of disagreement. Rather, it is an invitation for Zagzebski to explain where her view lies on the spectrum between views that accord disagreement a great deal of epistemic weight in all cases, and those that don't. Of greater interest, however, is the substantive question, *How do we rationally resolve the problem of disagreement?* As we've already seen, Zagzebski's solution relies on the notion of conscientious self-reflective trust. In fact, this notion plays a prominent role throughout her book. Thus, in the remainder, we'll consider it in some detail.

#### DISAGREEMENT AND CONSCIENTIOUS SELF-REFLECTIVE TRUST

It will help to begin with Zagzebski's general picture of rationality. As self-conscious beings who desire truth, we are in the unfortunate predicament that there is no non-circular way to determine that our belief forming faculties and practices on the whole are reliable. Given this limitation, we must simply trust that truth is attainable and that our faculties and environment are suitable for its attainment. However, even given such trust, we sometimes recognize that our cognitive faculties are not providing stable or consistent deliverances. This conflict between our mental states will often – though not always – result in an experience of dissonance. And often our executive self will pre-reflectively attempt to resolve dissonance by removing or adjusting one or many of the items responsible for generating that experience. Zagzebski's notion of rationality is derived from this natural practice of resolving dissonance. For Zagzebski, rationality is just the 'property we have when we do what we do naturally, only we do a better job of it. To be rational is to do a better job of what we do in any case – what our faculties do naturally' (p. 30). Since we naturally resolve dissonance, our rationality is proportionate to how much we improve on this natural practice.

Not all dissonance is resolved automatically and without reflection. In addition to trusting our cognitive faculties, we trust that being deliberative and careful in using our cognitive faculties – that is, being conscientious – is the best way to satisfy our desire for truth. Our trust in reasons or evidence is ultimately derived from our more basic trust in

epistemic conscientiousness, since searching for evidence or reasons is something that a fully conscientious agent does. Zagzebski states:

A conscientious person has evidence that she is more likely to get the truth when she is conscientious, but she trusts evidence in virtue of her trust in herself when she is conscientious, not conversely. Her trust in herself is more basic than her trust in evidence, and that includes evidence of reliability. The identification of evidence, the identification of the way to handle and evaluate evidence, and the resolution of conflicting evidence all depend upon the more basic property of epistemic conscientiousness. (p. 49)

When we experience dissonance and are unable to resolve it pre-reflectively, the rational response is to exercise epistemic conscientiousness in order to determine what is true. Interestingly, exercising epistemic conscientiousness does not just involve deliberate and careful consideration of *p*. For Zagzebski, the most rational response to dissonance involves consideration about the future and whether one's belief in *p* will survive 'future conscientious reflection, including reflection on future experiences, and future judgments about the past and present. The role of future conscientious self-reflection means that there is an important way in which the future justifies the present ...' (p. 50). In other words, one's belief that *p* is rational not because one has *judged p* to be true on the basis of conscientious self-reflection, but rather because one *now conscientiously judges that one will conscientiously judge p to be true*.

Disagreement produces dissonance within my deliberative reasons because though I trust my cognitive faculties and my conscientious use of my faculties, when I am conscientious I come to believe the following:

- (1) *p*.
- (2) Others have the same property (i.e. conscientiousness) that I trust in myself and believe not-*p*.
- (3) This similarity to me gives me *prima facie* reason to trust them.
- (4) Given my conscientiousness I have *prima facie* reason to trust myself.

The rational and conscientious agent, therefore, has to decide how to resolve dissonance by appeal to what she conscientiously believes will survive future conscientious self-reflection:

Disagreement with people we conscientiously judge to be conscientious should be handled ... in a way that we conscientiously judge will survive conscientious self-reflection. ... In some cases the disagreement will eventually be settled by evidence both parties accept, but that is not the case for all disagreements, and in any case, what is relevant is not whether the disagreement will be settled by future or present evidence that somebody has somewhere. What is relevant for me is what I conscientiously believe, and what I predict will satisfy my future self-reflection, given what I conscientiously predict about myself. (p. 215)

This is further evidence that Zagzebski's view offers a kind of middle-way between the 'extreme positions' she describes. For on the face of it, in some cases I might conscientiously judge that my belief is most likely to survive future conscientious reflection if I trust my dissenter more than myself; in other cases, I might judge that trusting myself is the most likely to yield a belief that will survive future conscientious reflection; and in others still I might suspend judgment about this matter – and therefore suspend about the target proposition. This entails that neither egoism nor egalitarianism yields the correct result in all cases of disagreement. It also implies a result that defenders of the total evidence view have been eager to emphasize: there's no single correct solution to the problem of disagreement; that is, there's no single attitude (belief, disbelief, suspension of judgment) that is rational to hold in all cases of disagreement. Rather, what is rational to think will depend on the details of a given case.<sup>5</sup>

However, we are also now in a position to see how Zagzebski's view is *different* from the total evidence view and that her solution is extreme in its own way. The following close paraphrase reconstructs how she recommends dealing with the dissonance generated by the numbered propositions above:

Awareness of this set of beliefs produces psychic dissonance and the desire to resolve the conflict by giving up one of the beliefs. [To] that end I will ask myself which beliefs are more likely to satisfy my future self-reflection. If upon reflection I trust (2) and (3) more than I trust (1) and (4), it is reasonable for me to adopt a skeptical attitude towards (1). ... But there are many other possibilities. I might trust (1) and (4) more than (2), in which case I would have no reason to give up (1) even

---

<sup>5</sup> See Thomas Kelly, 'Disagreement and Higher-Order Evidence', in Richard Feldman and Ted A. Warfield (eds.), *Disagreement* (New York: OUP, 2010) on this point.

if I have a high degree of trust in (3). Alternatively, I might trust (1) and (4) more than (3). I have given reasons for adopting some form of the principle of trust in earlier chapters, but there is no reason to assume that my trust in that principle will exceed my trust in every other belief I possess. I have reasons for accepting (4) and may continue to accept it upon conscientious reflection, but maybe I can say the same thing about my belief (1).<sup>6</sup>

It is telling that in this passage the different possible rational responses are all within one case and within one subject. In other words, Zagzebski seems to not only allow that in some cases it is rational to retain one's initial belief while in other cases it is rational to suspend judgment, but also that for any particular case it can be rational to retain one's initial belief and it can be rational to abandon that belief. Rationality boils down to what I conscientiously trust will survive my future conscientious reflection, and there is no way of saying in advance where that trust will fall. In this way, then, her view seems much more closely aligned with an egoist position. The character of the evidence, independent of what the conscientious agent takes it to be, drops out of the picture in determining whether one is rationally handling a particular case of disagreement. On her view, however, what matters is not what I *have* conscientiously judged to be true prior to the disagreement but on what I *now* conscientiously judge to be most likely to survive my *future* conscientious reflection.<sup>7</sup>

#### CONSCIENTIOUS SELF-REFLECTIVE TRUST: GENERAL CONSIDERATIONS

What is distinctive about Zagzebski's view is the role she assigns to conscientious self-reflection, and more specifically to future conscientious self-reflection. On her view, such conscientious judgment about the deliverances of future self-reflection seems to be necessary and sufficient not only for rational belief in the face of disagreement, but also

---

<sup>6</sup> See pp. 216-17 for the exact quote. We have changed the numbering in this passage to reflect the numbered propositions above. Zagzebski gives a very specific example of disagreement involving eight propositions. For ease of exposition we have attempted to provide a general formula that captures the essence of Zagzebski's more specific example.

<sup>7</sup> The higher-level requirement here is explicit: 'Disagreement with people we conscientiously judge to be conscientious should be handled ... in a way that we conscientiously judge will survive conscientious self-reflection.' (p. 215)



for rational belief in any case.<sup>8</sup> However, there are reasons for doubting both of these claims, and additional reasons for doubting whether the appeal to our future selves is likely to provide the epistemic guidance we seek. But if Zagzebski's *general* views about the role of conscientious self-reflection are in doubt, then her views about disagreement are also insecure. Accordingly, we'll now consider the necessity, sufficiency, and helpfulness of Zagzebski's brand of self-reflection vis-a-vis rational belief.

Let's start with *necessity*. Is it necessary for rational belief that *p* that I now conscientiously judge that *p* would survive future conscientious self-reflection? Plausibly not. Consider a physicist working to confirm or

---

<sup>8</sup> In this section, we attribute to Zagzebski the view that conscientious reflection on one's future conscientious self-reflection is both necessary and sufficient for rational belief of the sort that interests her. We think that this is a fair interpretation of her view. However, we admit that it is not perfectly clear whether Zagzebski consistently means to invoke this higher-level requirement. In some passages, she seems to slide between first-level and higher-level requirements on rationality. For instance, on p. 50 she states: 'The line of reasoning of this chapter has the consequence that ultimately our only test that a belief is true is that it survives future conscientious reflection, including reflection on future experiences, and future judgments about the past and present.' Note that this sentence begins with the requirement that a belief *in fact* survive future conscientious reflection (a first-level claim), while the clause 'including reflection on future experiences' moves the discussion one level up. In light of other passages (e.g., the passage quoted above from p. 215, which is intended as an application of her general view to a specific case), we judge that on balance, Zagzebski favours the higher-level requirement we attribute to her. But because there is room for dispute here, we intend this section to accomplish the following *disjunctive* task: it either shows that Zagzebski's view is problematic on account of its higher-level requirement *or* that her view should be further clarified so as to explicitly reject such a higher-level requirement.

We hasten to add that merely renouncing the higher-level view may not solve all the problems that lurk nearby. Among these: suppose Zagzebski's official view is that rational beliefs must *in fact* survive future conscientious reflection, and that this is also sufficient for rationality. In that case, she will need to explain why surviving future conscientious reflection is necessary and sufficient (or even helpful) for obtaining rational beliefs, given that we're not often aware of what such reflection recommends. On the necessity point, it seems plausible that a belief can be made rational by a subject's evidence at time  $t_1$ , even if that belief would not survive conscientious self-reflection at  $t_2$  (say, because the subject has new defeating evidence then). On sufficiency, it seems that even a belief that survives conscientious self-reflection at  $t_2$  may not have been rational at  $t_1$  (perhaps at  $t_1$  the subject believes for bad reasons, but comes to believe upon better reasons later). Why think that reflection at  $t_2$  justifies the subject's belief at  $t_1$ , especially if at  $t_1$  the subject is not aware of the former? Finally, it is difficult to see how future conscientious self-reflection can provide present epistemic help, given that we are not often aware of the conscientious reflections our future selves will undertake. Thanks to an anonymous referee for pressing us to consider this 'first-level' interpretation of Zagzebski's view.

disconfirm the existence of the Higgs Boson. She designs her experiments in painstaking detail, and carries out the experiments with similar care. She repeats the experiments many times, and assesses the relevant data as conscientiously as anyone could. Finally, she forms the belief that the Higgs Boson exists. However, at no point in this process does she consider whether her belief in the particle will survive conscientious self-reflection on the part of her future self. Her belief seems perfectly rational, but Zagzebski's view seems incompatible with this judgment.

Worse still, Zagzebski's requirement on rationality appears to entail that few of our beliefs are rational. When reflecting on whether to believe some proposition,  $p$ , we usually only consider the first-order evidence indicating whether  $p$  is true or false. In our more careful and reflective moments we may also consider higher-order evidence regarding  $p$  (e.g. whether other reliable or unreliable people with regards to  $p$  believe  $p$ ). But rarely – and for some of us probably never – when considering whether to believe  $p$  do we think about what our future conscientious self will also believe  $p$ . It follows, on Zagzebski's requirement, that few of us have very many rational beliefs.

Crucially, these problems arise not just when we compare Zagzebski's requirement to our everyday conception of rationality (though that would be problematic enough). Rather, the problems arise even on Zagzebski's comparatively more intellectualist view of rationality. On that conception, rationality consists in doing better what we do naturally. But as the above considerations show, we can do better – perhaps *much* better – than we do naturally, without making judgments about what conscientious judgments our future selves will make. For example:

- We might spend a great deal of time carefully and conscientiously evaluating whether our first-order evidence supports a scientific hypothesis;
- We might carefully and conscientiously consult both first and higher-order evidence in determining what to believe in a case of disagreement;
- We might devote an entire career to seeking and evaluating evidence relevant to God's existence.

Suppose that after such inquiries, we make judgments about the corresponding target propositions. On most evaluations, the resulting beliefs will be strong candidates for the post of *rational* belief. Of course, we sometimes make performance errors in evaluating our evidence,

so there's no guarantee of true or even rational belief at the end of our inquiries. But because the inquiries described above do not involve the very specific sort of self-reflection Zagzebski has in mind, the resulting beliefs are not even *candidates* for the accolade of rationality. No matter how carefully they are formed, if they don't result from conscientious reflection on the conscientious reflection of one's future self, these beliefs fail to be rational in Zagzebski's sense.

Now consider whether it is *sufficient* for rational belief in  $p$  that I currently conscientiously judge that I would judge  $p$  to be true on the basis of conscientious self-reflection. Here again, a negative verdict seems in order. Recall our scientist. Suppose that at time  $t_1$ , in the midst of her research, she considers what her future self will believe at some later time  $t_2$  about the Higgs Boson, once this future self engages in conscientious self-reflection about the matter. At  $t_1$ , she conscientiously judges that her future self will believe in the particle. She concludes that her future self will indeed hold this belief, and thus forms the belief herself. However, despite her conscientiousness at  $t_1$ , she badly misjudges her evidence relevant to the claim that

*Future*: My future self will believe in the particle after conscientious self-reflection.

Suppose that the weight of her current evidence actually supports the negation of *Future*, but that she ends up believing *Future* due to a cognitive malfunction that causes an egregious performance error. If this happens, then it seems her belief that *Future* is true is irrational, and consequently, that her belief about the Higgs Boson is irrational (at least if *Future* is its sole basis). Supposing this is right, we now have a case in which someone's current conscientious judgment about her future self's conscientious belief is insufficient to render that belief rational.

In addition to being neither necessary nor sufficient for rational belief, in many cases such higher-order reflection and future speculation is epistemically *unhelpful*. Consider again the scientist. She is hard at work in the lab and carefully evaluating the evidence for the existence of the Higgs Boson. It seems that turning her attention from the first-order evidence to what her future self is likely to believe upon future consideration is unhelpful with respect to her goal of believing the truth. If her colleagues catch her in the midst of such reflection, they will rightly exhort her to stop procrastinating and to get back to work. For surely what's most important in determining whether the particle

exists is the experimental data itself. If a judgment is to be made about the Boson, it should be made primarily on the basis of such evidence. If she indulges the distraction about her future escapades in conscientious self-reflection, our scientist merits a scolding. Reflection on her future self isn't making her a better scientist.

We can expand on this point by noting additional complications that may arise in the course of current conscientious reflection on the judgments of one's future self. First, it is unclear *which* future self should be the subject of one's reflection. Of course, it is absurd to suggest some uniform *temporal* criterion here. It is not as though what's relevant will always be reflection on what one's next-Tuesday-self or one's next-St.-Patrick's-Day-self will judge to be true. But if not these days, *when*?

We suggest that any plausible answer will feature not some specific temporal event, but rather, expectations about one's future *reliability*. Presumably, in consulting the conscientious reflections of one's future self, one aims to gain epistemic leverage from a source of information that is better than one's current self. This suggestion provides a way to clarify the epistemic relevance of reflection on one's future self.

But even here, problems lurk. Consider our scientist again. Either her future self (whichever one she happens to focus on) will be less reliable, equally reliable, or more reliable with respect to  $p$  than she is now. If she has reason to think her future conscientious self is *less reliable* with respect to  $p$  (e.g. perhaps she will be the unfortunate victim of some mentally disabling disease), then there is no good reason to now trust her future judgment, even if will be conscientiously formed, instead of her *current* judgment. If her future self is *equally reliable* regarding  $p$ , then appealing to her future self, at most provides a present-future case of peer agreement or peer disagreement. However, there is still no reason for her to defer to her future judgment. Rather, it seems to us that she should just reflect carefully on the current first-order and higher-order evidence (which will in this case involve evidence about her future self) regarding  $p$  and trust her current conscientious judgment regarding  $p$ . Therefore, only if she has good reason to think that her future self will be *more reliable* than her current self-regarding  $p$  would it be appropriate to defer to the authority of what her future self would conscientiously believe.

Even supposing she has good reasons to believe that her future self will be either equally or more reliable than she is now (not a small supposition since it seems that we are often without the requisite evidence to make

such judgments with sufficient reliability), often she will not be privy to what her future self will in fact believe with respect to  $p$ . For example, if her future self will be more reliable with regards to  $p$ , presumably this will be because either she will have evidence about  $p$  that she does not currently possess or because she will be better able to see some logical relation between her current evidence and  $p$  that she cannot currently see (or both). But since she does not currently have that evidence or that insight she will often be unable to judge what attitude her future self will take with respect to  $p$ . Therefore, the only way it could be rational to appeal to her future conscientious self when considering whether to believe some proposition is if 1) she has good reason to believe that she will be more reliable with respect to that proposition and 2) she has good reasons concerning what attitude her future self will have toward that proposition. Unfortunately, it is far from clear that these conditions are often satisfied.

Turning our attention back to the problem of disagreement, it seems to us that Zagzebski's solution is less than compelling for the reasons that her more general account of rationality is problematic. It seems to us that her almost exclusive focus on conscientious self-reflection comes at the cost of ignoring two important factors for rational belief: the character of our first-order and higher-order evidence, and how reliable we are in handling that evidence. Zagzebski's account seems incomplete to the extent that she ignores these items. For our part, we're confident that our future selves will have more information about the soundness of the above critiques once we have read Zagzebski's replies to them. In at least that respect, our future selves will be in a better epistemic position than our current selves. However, we are unsure what evidence Zagzebski will supply in responding to our arguments. Thus, reflecting on the conscientious reflection of our future selves is currently of little help to us – a fact which both gives us pause about the soundness of our criticisms and leaves us eagerly awaiting Zagzebski's replies.<sup>9</sup>

---

<sup>9</sup> Thanks to Tomás Bogardus and an anonymous referee for helpful comments and discussion.