

RESEARCH ARTICLE

Transformative experience, awareness growth and the limits of rational planning

Katie Steele*¹ and H. Orri Stefánsson^{2, 3}

¹ School of Philosophy, Australian National University, Canberra, Acton ACT 2601, Australia

*Corresponding author. Email: katie.steele@anu.edu.au

² Department of Philosophy, Stockholm University, SE-106 91 Stockholm, Sweden

Email: orri.stefansson@philosophy.su.se

³ For helpful feedback, thanks to our fellow symposium members and audience at the PSA Meeting in November 2021, to the audience of the ANU Humanising Machine Intelligence seminar, and to an anonymous reviewer for this journal. Steele is grateful for support from an ANU Futures Grant. Stefánsson gratefully acknowledges financial support from Riksbankens Jubileumsfond.

(Received XX XX XXXX; revised XX XX XXXX; accepted XX XX XXXX)

Abstract

Laurie Paul (2014, 2015) argues that, when it comes to many of your most significant life-changing decisions, the principles of rational choice are silent. That is because, in these cases, you anticipate that one of your choice options would yield a *transformative experience*. We argue that the best way to make sense of transformative decisions is to see them as ones in which you anticipate *awareness growth*. You do not merely lack knowledge about *which* possible outcome will arise from a transformative option; you lack knowledge about *what are* the possible outcomes. We indicate how principles of rational choice can be extended to cases of anticipated awareness growth.

1. Introduction

Laurie Paul's extremely influential work (2014, 2015) raises an important question: What are the *limits of rational planning*? Paul focuses on what she calls "transformative experiences", which are, roughly speaking, big changes or discontinuities in how one understands the world and evaluates states of affairs, like when a blind person gains sight or a formerly childless person becomes a parent. According to Paul, such changes are *epistemically transformative*, in that the person gains epistemic access to a fact about personal experience in virtue of having the experience in question. This, Paul claims, puts limits on rational decision-making (and all the more so when epistemic transformations are accompanied by changes in preferences). There will be many choice situations, she claims, in which a person predicts that one of her options will yield an epistemically transformative experience that is moreover relevant to the evaluation of this option. The person's current epistemic state, which serves as the basis for her current evaluation of the options, is thus relatively impoverished.

Critics of Paul's arguments have taken issue with various of her claims about what makes transformative experiences extraordinary for decision-making (e.g., Dougherty et

al. 2015, Harman 2015, Bykvist and Stefánsson 2017). Here, however, we will grant the following parts of Paul’s account. In the example of having a child, we will assume that there are some agents for whom their personal phenomenological experience matters for the evaluation of whether or not to have a child, and think that only by having a child can they get epistemic access to aspects of phenomenological experience that are important for evaluating this course of action.

Nonetheless, important questions remain. Is the shift in epistemic access in Paul’s examples really so radical as to warrant the label “epistemically transformative”? And are the evaluations of options that the agent is able to make prior to the experience in question so ill-informed, by her own lights, as to make rational decision-making impossible? Here we answer “yes” to the first question, in line with Paul, but “no” to the second, parting ways with Paul. We argue that the shift in epistemic access in Paul’s examples can be seen as radical if treated as an episode of *awareness growth*, which means an augmentation in the agent’s “possibility space”, or what she takes to be all the different ways that the world might be. But recognition of one’s current *limited awareness*, perhaps because one anticipates later awareness growth (as in the examples here), does not make rational choice impossible.

In Section 1, we present a model of what we call “anticipated awareness growth” and make the case that the kind of decision-making predicament Paul describes is best seen as transformative in just this way. However, we argue in Section 2 that this does not preclude making rational decisions in (more or less) the usual way. We propose a principle of coherence that reveals how a decision-maker may go about evaluating outcomes that she cannot yet specify in detail. In Section 3 we compare our account with an alternative and defend it against potential objections.

2. Transformation as Awareness Growth

Paul picks out phenomenological facts—what it feels like to have certain kinds of experiences—as being beyond one’s epistemic access in some unusual sense, prior to actually having that kind of experience. Paul starts from the idea that all the knowledge that one can amass about third-person matters does not yield knowledge of some first-person matters—what it feels like to occupy certain positions in the world. She extends the lesson of Jackson’s (1986) Mary case to other sorts of evidence for first-person facts. According to Paul (e.g., 2015, p. 154), it is hard to think of any good source of evidence that Mary might draw on in her black-and-white room—including, say, others’ testimony about what it is like to see colour—that would allow her to fully appreciate this entirely new kind of experience.

Even if Paul underestimates testimony as a source of evidence for phenomenological experience, we think that her examples are in fact paradigms of some unusual kind of epistemic limitation that can be relevant for decision-making. The way to make sense of this, we contest, is to see the epistemic transformations as episodes of awareness growth. We defined this above as an augmentation in the agent’s possibility space. The agent comes to see that there are practically important ways that the world might be that she hadn’t previously considered. Such augmentations are the result of coming to better appreciate a property of the world or random variable: it may be that one recognises a further value that some property/variable might take (what is called awareness growth by *expansion*) or it may be that one recognises an entirely new property/variable (what is

called awareness growth by *refinement*). In the case of parenthood, you might experience awareness growth by expansion if you come to realise that there is a further state of fatigue that you had not recognised—a kind of tiredness that nonetheless allows one to function reasonably well—that matters in accounting for the costs of lack of sleep. On the other hand, you might experience awareness growth by refinement if you come to realise that your child’s level of sociability matters to how much you enjoy, as a parent, spending time in the company of others.

Why do we claim that Paul’s epistemic transformations are best understood as awareness growth? The best argument for this claim, we suggest, is a negative one: it is hard to see that there are any better alternatives for understanding epistemic transformation and the preceding epistemic limitations. Paul speaks of a lack of *knowledge* (e.g., 2015, pp. 155–158). Is it simply that the agent initially lacks knowledge of *which* possible phenomenological experience will occur? That kind of epistemic shift is surely not sufficiently remarkable. An initial lack of knowledge about some decision-relevant matter, due to, say, not believing or having low confidence in the truth of the matter, is an entirely ordinary predicament and the paradigm case for standard decision theory. One could turn to more radical, non-standard kinds of uncertainty, such as a temporary inability to specify precise probabilities for the relevant possibilities.¹ But even this kind of epistemic predicament does not seem sufficiently extraordinary to capture the cases at hand. An inability to even specify or grasp the relevant possibilities is a much better candidate for spelling out this unusual lack of epistemic access. One might put it this way: it is not that the agent lacks knowledge of which of some possible phenomenological experiences she will experience upon taking some course of action. Rather, she lacks knowledge of what *are* the possible experiences associated with the course of action.²

Admittedly, limited awareness and subsequent awareness growth is not limited to personal phenomenological experience. Any kind of novel course of action, whether the novelty be due to the phenomenology of the possible outcomes or other “third-person” properties of the outcomes, can be characterised by limited awareness. But we do not see why this lack of exclusivity should be a problem. We thus proceed to model the decision-problems of interest as ones in which the agent recognises herself to have limited awareness, partly in view of the fact that she anticipates that some of her options would lead to awareness growth. Hence we will refer to the decision-maker’s predicament as one of *anticipated awareness growth*.

Let us continue with the example of the agent deciding whether to become a parent. It is useful to first consider a decision-problem where there is *no* anticipated awareness growth, as per table 1. The agent assumes there are n states of the world (from s_1 to s_n), that determine the outcome of the available options, which we assume here to be simply “become parent” (p) and “remain childless” (c). We will assume that, in addition to her uncertainty about other relevant matters, the states capture the agent’s uncertainty about the phenomenological experience of having a child and of not having a child. Finally, $p(s_i)$ is the outcome that obtains when p is chosen and s_i is the true state of the world.

¹This is how Dougherty et al. (2015) and Collins (2015) suggest we make sense of the epistemic limitations prior to a transformative experience.

²Note that Carr (2015) and Bykvist (2019) also suggest that epistemically transformative experience may be understood as awareness growth, but neither offers a model for practical decision in such circumstances, as we do in this paper.

	s_1	...	s_n
become parent (p)	$p(s_1)$...	$p(s_n)$
continue childless (c)	$c(s_1)$...	$c(s_n)$

Table 1. Prospective parenthood without anticipated awareness growth

What does it mean to say that table 1 does not make room for awareness growth? It means that the states s_1 to s_n are thought, by the agent,³ to exhaust the set of possible contingencies. Moreover, the outcomes that these contingencies yield under each option, $p(s_i)$ and $c(s_i)$, are thought to be maximally specific in all ways that might be relevant to the decision.

	E_1	...	E_n	??
become parent (p)	$p(E_1)$...	$p(E_n)$	$p(??)$
continue childless (c)	$c(E_1)$...	$c(E_n)$	$c(??)$

Table 2. Prospective parenthood with anticipated awareness growth

Consider now the matrix in table 2. We focus initially on the anticipation of awareness growth by *refinement* that is represented here. Suppose that E_1 to E_n is the agent’s finest and most exhaustive “partition” of the space of contingencies. (We see below that it is not really a partition, according to the agent.) Each E_i is not taken to be a (fully specified) state of the world, but rather an *event* that amounts to a set of states. It follows that the outcomes— $p(E_i)$, $c(E_i)$ —are not thought to be fully specified in all ways that matter to the decision. Perhaps the agent recognises that there is some aspect of the experience of having a child that she thinks matters to her evaluation, but she cannot really specify what this aspect of experience is, let alone the different ways it may be realised. The agent recognises her limited awareness, let us assume, but does not know how to further partition the E_i into the states that fully determine the outcomes of the option in all relevant detail.

Now we turn to the anticipation of awareness growth by *expansion* that is represented in table 2. The “??” column in the table represents that there is some further event that the agent cannot specify and yet considers part of the event space, such that the other events, E_1 to E_n , are not an *exhaustive* set of mutually exclusive events. We call the “??” event a “subjective catch-all” (Steele and Stefánsson 2021b: sec. 6.3). In other words, the agent is aware of the fact that there might be contingencies that are inconsistent with all of E_1 to E_n and which would, if they materialise, determine the outcome of the two available options.

Note that, formally speaking, both types of anticipated awareness growth are cases of refinement. While we call the latter case “expansion”, it is modelled as if the agent anticipates *refining* her catch-all. Indeed, if an agent anticipates awareness growth of any kind, then she must already account for it in some way, be that via crudely described outcomes, the most extreme being a subjective catch-all. This is further reason why our account lines up well with Paul’s characterisation of decision-problems that have the

³Recall that what is of interest is a decision situation where an agent suspects that an option might bring about transformative experience, which we want to analyse as awareness growth. To make sense of this, we need to assume that the lack of awareness we are about to model is a lack *from the agent’s perspective*. See Steele and Stefánsson (2021b: sec. 3) for a further discussion of modelling limited awareness from an agent’s perspective vs. from a modeller’s perspective.

potential to bring about transformative experience. For Paul, the crucial limitation is not that some possible outcomes are not accounted for at all, but rather that their values are unknown. Our model is not inconsistent with this basic idea. But it will be shown in the next section that the details of anticipated awareness growth matter; while the agent does not know for sure the value of some outcomes, she may nonetheless assign them an expected value.

3. Managing Awareness Growth

The question now is whether anticipated awareness growth presents a special obstacle to decision-making. Does the shift from a decision-problem like that in table 1 to one like that in table 2 require a radical shift in the notion of rational choice? We will argue that it does not, and in this sense, our account of transformative decisions diverges from that of Paul. Paul's position is that one cannot choose rationally in situations where the options may result in a transformative experience. Here we argue that this is too pessimistic—agents can reason about the value and probability of catch-all and other not-fully-specified outcomes in a relatively ordinary way. In the next section, we respond to objections and compare our view with a rival approach.

We offer a way to reason about the value of the catch-all. (We focus on anticipated “expansion”, to simplify our analysis. But all of what we say applies to anticipated “refinement” too.) The first thing to note is that there is general agreement amongst scholars working on (anticipated) awareness growth that it is possible to infer from an agent's preferences over options what is the value they assign to a catch-all. We can examine an agent's conditional preferences. Suppose that one prefers becoming a parent to remaining childless, conditional on events E_1 to E_n . Still, overall one prefers to remain childless. Is this inconsistent? No, it simply reveals that one prefers remaining childless to becoming a parent conditional on the catch-all, and sufficiently so that this overturns the preference for becoming a parent conditional on the other events (in light of the probabilities for these events).⁴

But Paul is unlikely to be impressed by this appeal to a “preference-first” (Pettigrew 2017) conception of decision theory and catch-all evaluation. Paul is interested here in a “deliberative” conception of decision theory (to continue with Pettigrew's terminology), whereby the agent appeals to decision theory to arrive at well-reasoned preferences. We want to know not simply *that* the agent assigned some value to outcomes under a catch-all event that yielded consistent overall preferences, but *why*, or *on what basis*, she assigned those values and thus had those preferences.

What we propose is a coherence principle that indirectly bears on deliberation in that it suggests a procedure for arriving at the values of outcomes associated with a catch-all. We say more on the deliberative procedure below. First, let us state the coherence principle, which we dub *Awareness Reflection (outcome values)*. It concerns how the agent's current evaluation of an outcome relates to her predicted future evaluations of that outcome (in so far as she makes such a prediction and considers these future attitudes to be more enlightened) when she anticipates awareness growth. As the name suggests, the principle is inspired by a similar principle known simply as *Reflection* and defended by

⁴Piermont (2017) calls such conditional preferences “contingency plans”, and uses them to formally characterize anticipated awareness growth (or “introspective unawareness”, to use his term); see too Karni and Vierø's (2017) representation theorem for anticipated awareness growth.

van Fraassen (1984), concerning probabilistic belief and applicable in ordinary situations that do not involve anticipated awareness growth.⁵ In short, the reflection principle that we will explore below states that an agent's evaluation of an outcome, conditional on having some evaluation given awareness growth (and no other personal change), should just be that more aware evaluation.⁶

We now state the core principle more formally. Let v and v^+ represent cardinal value functions that are comparable in units and levels. Assume that v represents the agent's current value function. Let \mathbf{v}^+ stand for the proposition that the agent will have value function v^+ at time t , having experienced awareness growth and no other personal change that would mean she was less enlightened. Then Awareness Reflection (outcome values) can be stated as follows:

Awareness Reflection (outcome values). For all options f and all events E_i (including the catch-all), and for any v and v^+ , and for all \mathbf{v}^+ of which the agent is aware:

$$v(f(E_i) \mid \mathbf{v}^+) = v^+(f(E_i))$$

The agent's current value for any catch-all outcome is thus constrained by her predicted future values for this outcome.

In special circumstances, the agent's current value for any given outcome is the expectation of her future values for this outcome. The special circumstances are those in which the following conditions are met. First, the agent is sure she will experience awareness growth and no other personal change. Moreover, she predicts that her probability distribution over the event space, E_1, \dots, E_n , plus the catch-all, $??$, will be constant upon awareness growth (more on this assumption below). So she predicts that only her values will change upon awareness growth. The various candidates for her future value function form a partition of her possibility space, represented by $\{\mathbf{v}_1^+, \dots, \mathbf{v}_z^+\}$.⁷ Furthermore, the agent is indifferent to (i.e., neither desires nor detests) whatever values she will have once her awareness grows (since she assumes these will be the result of only having undergone awareness growth). The latter means that we can assume that $v(f(E_i) \mid \mathbf{v}^+) = v(f(E_i) \& \mathbf{v}^+)$.⁸ Now, since $\{\mathbf{v}_1^+, \dots, \mathbf{v}_z^+\}$ is a partition of the agent's possibility set, Jeffrey's (1965) conditional expected value formula implies that:

$$v(f(E_i)) = \sum_{i=1}^z P(\mathbf{v}_i^+ \mid f(E_i)) \cdot v(f(E_i) \& \mathbf{v}_i^+)$$

⁵Arntzenius (2008) extends van Fraassen's principle to desire.

⁶As we discuss in Steele and Stefánsson (2021b: sec. 7.3.2), the principle should strictly speaking be weakened to saying that one's current conditional evaluation should neither be more nor less positive than the more aware evaluation, which is consistent with the former being undefined even though the latter is not. We set aside this complication here.

⁷More generally, we might say that an agent's current value function should match her expected future value function, *conditional on awareness growth and no other personal change* and assuming her probabilities over the event space are not predicted to change in these circumstances.

⁸In fact, Joyce (1999) assumes that the equality $v(f(E_i) \mid \mathbf{v}^+) = v(f(E_i) \& \mathbf{v}^+)$ always holds. By contrast, Bradley (1999) argues that $v(f(E_i) \mid \mathbf{v}^+) = v(f(E_i) \& \mathbf{v}^+) - v(\mathbf{v}^+)$, which collapses into $v(f(E_i) \mid \mathbf{v}^+) = v(f(E_i) \& \mathbf{v}^+)$ in the above-mentioned 'special circumstances', since then $v(\mathbf{v}^+) = 0$ (by the convention of reserving 0 for that which the agent neither desires to be true nor false).

But then Awareness Reflection (outcome values) and the above conditions together imply that:

$$v(f(E_i)) = \sum_{i=1}^z P(\mathbf{v}_i^+) \cdot v_i^+(f(E_i)) \quad (1)$$

One of the above assumptions is that the agent predicts her probabilities over the event space to remain constant upon awareness growth. Otherwise the current value of outcomes must be sensitive to how future values and future probabilities are predicted to co-vary. This more complicated case can be captured by a more general reflection principle concerning how current all-things-considered preferences should relate to predicted more aware all-things-considered preferences. See Steele and Stefánsson (2021b: sec. 7) for the formulation of such a principle, dubbed *Preference Awareness Reflection*. Its defence turns on the importance of avoiding sure loss in a sequential-choice setting (ibid). The specific reflection principle outlined here—Awareness Reflection (outcome values)—can be derived from the more general Preference Awareness Reflection. See the Appendix for further details.

Reflection principles are often thought implausible in requiring an agent to defer to her future self, even if she does not respect this future self in the appropriate sense. But note that Awareness Reflection (outcome values) is highly restrictive since it concerns only those future states of greater awareness that do not involve any other personal change that would make the agent more defective. Moreover, the agent must actually entertain these possibilities regarding her future values. So Awareness Reflection (outcome values) may rarely be applicable.

Let us finally return to how the principles described above bear on an agent's deliberations, and suggest a template for *explaining* how she arrives at her current values for the catch-all outcomes. We suggest that when reasoning about the value of an outcome, under the assumption that one's evaluation of it might be affected by an anticipated (potentially transformative) growth in awareness, the best that an agent can do is to ask herself how someone who has undergone the anticipated awareness growth, but otherwise has the same preferences and beliefs, would judge the outcome in question.⁹ In particular, she should *defer* to someone who is exactly like herself except being more aware. Often she will not know for sure how her more aware self would evaluate an outcome. But she may consider some possible future attitudes more and less probable, in which case the best she can do is to try to align her current evaluation with what she currently expects her more aware evaluation will be.

4. Responding to scepticism

Our coherence principles and associated deliberative procedure leave some questions unanswered. In particular, even if one could form some opinion—in part by observing similar but more enlightened others—about one's more-aware evaluation of a catch-all, is it really plausible that one could have a probability distribution over the possible evaluations? And do the principles apply in cases of *personally transformative experience*, whereby there is a shift in one's fundamental values?

⁹This assumes that the evaluation in question is based on how desirable the outcome is *for the person*, which is indeed what Paul is interested in too.

One might see only negative responses to these questions, and therefore agree with Paul that there is no rational basis for one's current evaluation of a catch-all outcome. Perhaps an entirely different way of evaluating a catch-all is called for, that is not about trying to guess one's attitude to this catch-all, but rather reflects one's generic attitude about encountering the unknown. Karni and Vierø (2017) make a suggestion like this, and it seems to be in line with Paul's remarks about basing potentially transformative decisions on the "revelatory" value of discovering the content of the catch-all and one's attitudes towards it.

We think this kind of move, in terms of valuing catch-all outcomes, is a far last resort. It amounts to a rather *ad hoc* narrowing of the content of catch-all outcomes to a very inward-looking property: something like "surprise about one's attitudes". Moreover, there is more to say in favour of deliberating about how one's future self will view catch-all outcomes. Start with the latter of the above-mentioned issues. There are two stances one may take towards a personally transformative awareness growth. The personal transformation may be predicted to yield a more enlightened self, to whom one has good reason to defer, or else the personal transformation may be predicted to yield an entirely different self to whom one has no reason to defer. Either way, one can defer *just* to those future versions of oneself that are more enlightened. One's current values should simply be one's expectations of one's future values conditional on an "orderly (or enlightening) awareness growth" (recall footnote 7). Surely one has some confidence in such an awareness growth, however little.

Moreover, we hold that there is further inductive basis for predictions about what will be one's future values having experienced awareness growth: similar past decision-problems. The inference that we have in mind, based on one's past experience of similar circumstances, has been proposed by Grant and Quiggin (e.g., 2013). Our approach might be seen as a finer-grained version of theirs. Grant and Quiggin focus on environmental intervention and appeal to something akin to "Murphy's law"¹⁰ in this context: "Sometimes unexpected things happen and the unexpected is always bad." We suggest that the reference classes of similar decision-problems may be narrower, and the inferences more nuanced. For instance, one might infer that "in career-changing decisions, sometimes unexpected things happen, and sometimes the unexpected is moderately good and sometimes moderately bad". More generally, it would seem reasonable to have different expectations about the (un)desirability of the unexpected depending on what type of decision is involved.

5. Conclusion

In conclusion, we think that although Paul is right in pointing out that the shifts in epistemic access that her examples illustrate are radical and not captured by ordinary rules for learning, we do *not* think that she is correct in suggesting that such shifts make rational choice impossible. We have argued that the shifts in question are instances of awareness growth, and we have moreover suggested how the anticipation of such growth can be built into slight modifications of standard decision models. Thus, even though

¹⁰"Anything that can go wrong will go wrong."

becoming a parent, say, may bring about experiences that are epistemically transformative, that does not mean that one cannot rationally decide to become (or not to become) a parent.

References

- [1] Arntzenius, Frank. 2008. “No regrets, or: Edith Piaf revamps decision theory”, *Erkenntnis* 68(2): 277–297.
- [2] Bradley, Richard. 1999. “Conditional Desirability”, *Theory and Decision* 47(1): 23–55.
- [3] Bykvist, Krister. 2019. “Paul’s Reconfiguration of Decision-Problems in the Light of Transformative Experiences”, *Rivista Internazionale di Filosofia e Psicologia* 10(3): 346–356.
- [4] Bykvist, Krister and Stefánsson, H. Orri. 2017. ‘Epistemic Transformation and Rational Choice’, *Economics and Philosophy*, 33: 125–138.
- [5] Carr, Jennifer. 2015. “Epistemic Expansions”, *Res Philosophica* 92(2): 217–236.
- [6] Collins, John. 2015. “Neophobia”, *Res Philosophica* 92(2): 283–300.
- [7] Dougherty, Tom, Horowitz, Sophie and Sliwa, Paulina. 2015. “Expecting the Unexpected”, *Res Philosophica* 92(2): 301–321.
- [8] Grant, Simon and Quiggin, John. 2013. “Inductive reasoning about unawareness”, *Economic Theory* 54(3): 717–755.
- [9] Harman, Elizabeth. 2015. “Transformative Experiences and Reliance on Moral Testimony”, *Res Philosophica* 92(2): 323–339.
- [10] Jackson, Frank. 1986. “What Mary Didn’t Know” *The Journal of Philosophy* 83(5): 291–295.
- [11] Jeffrey, Richard. 1965. *The Logic of Decision*. Chicago: University of Chicago Press.
- [12] Joyce, James M. 1999. *The Foundations of Causal Decision Theory* Cambridge: Cambridge University Press.
- [13] Karni, Ed and Vierø, Marie-Louise. 2017. “Awareness of unawareness: A theory of decision-making in the face of ignorance” *Journal of Economic Theory* 168: 301–325.
- [14] Paul, L. A. 2014. *Transformative Experience*. Oxford: Oxford University Press.
- [15] Paul, L. A. 2015. “What You Can’t Expect When You’re Expecting”, *Res Philosophica* 92(2): 149–170.
- [16] Piermont, Evan. 2017. “Introspective unawareness and observable choice”, *Games and Economic Behavior* 106(C): 134–152.
- [17] Pettigrew, Richard. 2015. “Transformative Experience and Decision Theory”, *Philosophy and Phenomenological Research* 91(3): 766–774.
- [18] Steele, Katie and Stefánsson, H. Orri. 2021a. “Belief Revision for Growing Awareness”, *Mind*, 130(520): 1207–1232.
- [19] Steele, Katie and Stefánsson, H. Orri. 2021b. *Beyond Uncertainty: Reasoning with Unknown Possibilities*. Cambridge, UK: Cambridge University Press.
- [20] van Fraassen, Bas C. 1984. “Belief and the will”, *Journal of Philosophy* 81(5): 235–256.

Appendix

We propose a very general reflection principle that we dub *Preference Awareness Reflection* (Steele and Stefánsson 2021b: sec. 7.5). It says that the agent’s current preference relation for any *prospects* (including options and outcomes) *a* and *b* should not differ from her predicted future preference relation for these prospects after her awareness has grown, assuming no other personal change takes place. The way to think about this is that the agent’s evaluation of *a* and *b* should equal her expected future evaluation for *a* and *b* after awareness growth and no other change.

Let us spell this latter idea out formally for the outcomes of a given act *f* under any event *E* (including the catch-all) i.e., for *f*(*E*). Let \mathbf{FV}_i^+ ($i = 1, \dots, z$) be propositions

stating the different possible future attitudes of the agent, having experienced awareness growth and no other personal change. Assume these form a partition of the possibility space. Any given proposition \mathbf{FV}_i^+ says that the agent experiences awareness growth and no other personal change and subsequently has a probabilistic belief function, p_i^+ , and a value function, v_i^+ over the outcomes. Let p and v be the agent's current probability and value function respectively. The idea is that the current contribution of an outcome to the expected value of the option (the product of the current value and probability of the outcome) is the expectation of the future possible contributions of the outcome to the expected value of the option:

$$v(f(E)) \cdot p(E) = \sum_{i=1}^z p(\mathbf{FV}_i^+) \cdot (v_i^+(f(E)) \cdot p_i^+(E))$$

We are concerned with the special case whereby the probability of the event in question is thought to be constant, no matter how awareness grows. Thus $p_i^+(E) = p(E)$ for any i . Then:

$$v(f(E)) = \sum_{j=1}^z p(\mathbf{FV}_j^+) \cdot v_j^+(f(E))$$

This matches the expression (1) for the current value of an outcome (in the stated special circumstances) that appears in Section 3.