Transformative experience and the shark problem

Tim Campbell¹ · Julia Mosquera¹

Published online: 31 January 2020
© The Author(s) 2020, corrected publication April 2020

Abstract In her ground-breaking and highly influential book *Transformative Experience*, L.A. Paul makes two claims: (1) one cannot evaluate and compare certain experiential outcomes (e.g. being a parent and being a non-parent) unless one can grasp what these outcomes are like; and (2) one *can* evaluate and compare certain intuitively horrible outcomes (e.g. being eaten alive by sharks) as bad and worse than certain other outcomes even if one cannot grasp what these intuitively horrible outcomes are like. We argue that the conjunction of these two claims leads to an implausible discontinuity in the evaluability of outcomes. One implication of positing such a discontinuity is that evaluative comparisons of outcomes will not be proportionally sensitive to variation in the underlying features of these outcomes. This puts pressure on Paul to abandon either (1) or (2). But (1) is central to her view and (2) is very hard to deny. We call this the Shark Problem.

Keywords L.A. Paul · Transformative experience · Rational choice · Subjective value · The shark problem · Spectrum

Motherhood is more than I ever imagined. It's more exciting and more terrifying; more rewarding and more draining; easier to figure out yet totally confusing; it's a daily dose of the brand new and the mundane.

—Stephanie Thomas, Mother.

I've never felt fear in my life like what I felt in the jaws of that white pointer. I went straight into its mouth front onwards.

—Eric Nerhus, Shark Attack Survivor.

Tim Campbell and Julia Mosquera have contributed equally to this work.

✉ Julia Mosquera
julia.mosquera@iffs.se

Tim Campbell
timothy.campbell@iffs.se

¹ Institute for Futures Studies, Holländargatan 13, Box 591, 101 31 Stockholm, Sweden

1 Parenting and shark attacks

Many people at some point in their lives will have an experience that is radically different from any that they have had before. Possible examples include the experiences of becoming a parent, acquiring a new sensory ability, and tasting durian for the first time. L.A. Paul calls these radically new experiences ‘transformative experiences’ (Paul 2014: 15–18).

The phenomenological content of a transformative experience is not merely unknown but *inaccessible*. To illustrate this, Paul considers Frank Jackson’s famous example of Mary The Scientist, who, while isolated in her black and white room, has never seen the colour red. Mary cannot grasp *what it is like* to see red, despite knowing all the relevant science (e.g. about wavelengths, etc.) without seeing red herself (Paul 2014: 9–10).

According to Paul, experiential outcomes (e.g. an outcome in which one sees red; an outcome in which one is a parent, etc.) have a kind of value, what she calls ‘subjective value’ (Paul 2014: 12), that cannot be grasped by those who are unfamiliar with the precise phenomenology of the experiences that occur in those outcomes. Although obscure, the notion of subjective value can be understood as referring to a particular kind of value attaching to the precise phenomenology of an experience, and this kind of value is not reducible to the value of the painfulness or the pleasurable of the experience.¹

Paul’s project concerns the evaluation, comparison, and choice between different possible outcomes with respect to subjective value (Paul 2014: 25). Hence, unless we explicitly say otherwise, when discussing the evaluation of outcomes we will assume that this evaluation is made with respect to subjective value, and we will omit the locution ‘with respect to subjective value’.

Paul accepts what we call

The Prior Experience Claim: One cannot evaluate and compare different experiential outcomes unless one can grasp what these outcomes are like, which one can do only if one has previously experienced outcomes of that kind. (Paul 2014: 2, 71–94)

Paul’s justification for the Prior Experience Claim is this: Generally speaking, one is rationally required to estimate the subjective value of an outcome by cognitively modelling that outcome and mentally projecting one’s self into it, so that one can have some idea of what it would be like to be the subject of whatever experiences occur in that outcome.²

However, according to Paul, one cannot cognitively model outcomes that contain experiences of a phenomenological type that one is unfamiliar with (Paul 2014: 109–110). In this case, as in the case of Mary, one lacks the phenomenal concepts necessary to complete the required imaginative task. In sum, since subjective value

¹ For further discussion on the notion of subjective value see Bykvist and Stefansson (2017) and Talbott (2016).

² See Paul 2014, Ch. 2, especially pp. 19–28.

attaches to the precise phenomenology of an experience, one cannot estimate the former without grasping the latter; but grasping the latter is impossible if one has not had experiences of its phenomenological kind.

Interestingly, however, Paul does not think that the Prior Experience Claim is *fully general*. She claims that for certain outcomes that are intuitively horrible, one can know that they are bad and worse than certain other outcomes even if one does not know what they are like. One of Paul's examples is an outcome in which one is eaten alive by sharks. She says

In cases like [being eaten by] sharks, we don't need to perform an assessment of the outcome by cognitively modelling what it would be like, because we know what the results would be: we know every outcome is bad, *whatever* it is like (Paul 2014: 128; cf. 27).

So, in addition to the Prior Experience Claim, Paul accepts what we call

The Shark Claim: One can evaluate and compare certain intuitively horrible outcomes (e.g. being eaten alive by sharks) as bad, and worse than certain other outcomes even if one cannot grasp what these intuitively horrible outcomes are like. (c.f. Paul 2014: 127; cf. 27)

Paul discusses other examples such as being hit by a bus and having your legs amputated without anesthesia (Paul 2014: 28, 127; 2015: 802–3). According to Paul, one does not need to cognitively model these outcomes; rather, one can simply know that they are bad no matter what they are like (Paul 2014: 25). The theoretical justification for recognizing these cases as exceptions to the Prior Experience Claim is not entirely clear. However, *that they are exceptions* is intuitively plausible. Nobody, we hope, would deny that one can know that an experience of being eaten alive by sharks, for example, is bad even if one is not familiar with the precise phenomenology of that experience.

When discussing the Prior Experience Claim, Paul has in mind only certain kinds of experiential outcomes, those that are relevant to most of us. She explicitly sets aside outlandish cases, such as considering whether to be eaten alive by sharks, “in which there is no need to deliberate by cognitively modelling in order to assess the subjective value of the relevant outcomes” (Paul 2014: 28).

Unfortunately, as we show in Sect. 2, restricting the application of the Prior Experience Claim in order to accommodate the Shark Claim gives rise to a theoretical problem. We call it *the Shark Problem*. In Sect. 3, we defend a key assumption underlying the argument presented in Sect. 2. In Sect. 4, we consider and reject a solution to the Shark Problem based on Paul's suggestion that one can evaluate outcomes on the basis of what she calls revelatory value (Paul 2014: 38). Section 5 is the conclusion.

2 The experiential range view

According to the Experiential Range View (ERV), there are two different kinds of experiential outcomes. We define *normal outcomes*, such as being a parent, as outcomes whose subjective values are crucial to their evaluation, and to which the

Prior Experience Claim applies. And we define *sharky outcomes*, such as being eaten alive by sharks, being hit by a bus, getting your legs amputated without anesthesia, etc., as outcomes that can be evaluated as bad and worse than normal outcomes, and to which the Prior Experience Claim does not apply.³

A fully developed version of ERV will include an explanation as to why the Prior Experience Claim either does or does not apply to a particular set of outcomes. We leave open what this explanation might be.⁴ For we think that on *any plausible version* of ERV, a normal outcome can be gradually transformed into a sharky outcome. To see this, imagine a spectrum of possible experiential outcomes ranging from an outcome with a very small amount of intense pain, to an outcome with a large amount of intense pain, such that one can move from the first outcome to the last outcome in a finite series of steps, each involving the addition of only one more second of intense pain. The outcomes in this spectrum (S) are depicted in Fig. 1 below.

Each outcome (O_1 , O_2 , etc.) has a duration of 30 days⁵ and is represented by one of the horizontal lines distributed vertically in Fig. 1. Each notch in one of these horizontal lines represents one second of intense pain, as well as the temporal location of this pain within the 30-day outcome in which it occurs. The number of notches increases as one moves downward from O_1 , since each outcome after O_1 contains one more second of intense pain than the immediately preceding outcome. For example, O_1 might be a brief stretch of a life as a parent that includes one second of intense pain; O_2 might be a brief stretch of an experientially different life that includes two seconds of intense pain, etc. We assume that the extra second of pain in each step does not alter the temporal distribution of the other seconds of pain.⁶ Each of the solid bars in $O_{500,000}$ and $O_{500,001}$ represents many seconds of intense pain closely grouped together in time.

Importantly, each outcome in S is experientially different enough from the others that a person that has had experiences in only one of these outcomes would be unable to grasp the experiences in the other outcomes; she would be unable to know what these other outcomes are like. Thus, while it is true that, for example, $O_{500,000}$ and $O_{500,001}$ are phenomenologically similar with respect to the pain they contain, they are phenomenologically dissimilar enough that they cannot be compared by an agent who is unfamiliar with the experiences that occur in them. The amount of pain in $O_{500,000}$ is less than $\frac{1}{4}$ of its total duration (7 days out of a total of 30). Therefore,

³ It is worth mentioning that if there are sharky outcomes, then there might also be “heavenly” outcomes—good experiential outcomes to which the Prior Experience Claim does not apply. One can grasp that such outcomes are good without knowing what they are like. For more on this see Sullivan (2018: 165–7).

⁴ One possible explanation is offered by Talbott (2016: 383). Talbott interprets Paul as endorsing the view that the information about normal outcomes that is available to the agent isn’t “personalized enough” to enable her to make accurate evaluations of those outcomes, whereas the same is not true about sharky outcomes.

⁵ This stipulation is arbitrary; the outcomes could be shorter or longer.

⁶ We acknowledge that the way in which the pain in the outcomes is distributed might make a difference to the phenomenology of the outcomes (i.e., two outcomes containing 4 days of pain each, may be phenomenologically different depending on whether the 4 days of pain are consecutive or spaced out, for example every Monday of the month). For the sake of simplicity, we leave this complication aside.

Fig. 1 Spectrum of experiential outcomes (S)

even though $O_{500,000}$ and $O_{500,001}$ contain similar amounts of the same type of extreme pain, most of the experiences in $O_{500,000}$ are radically different from most of the experiences in $O_{500,001}$, and from most of the experiences in the rest of the outcomes in S . A crucial assumption here is that the number of possible types of experience is sufficiently large to fill in S . We defend this assumption in Sect. 3.

For our present purposes, what is important is that while certain outcomes in S , for example O_1 and O_2 , contain an amount of pain that is consistent with their being normal outcomes, other outcomes in S , such as $O_{500,001}$ and each subsequent outcome, contain enough intense pain to be sharky.

If ERV is true, then there must be a *boundary* between the range of normal outcomes and the range of sharky outcomes. In our setup, this boundary is precise— $O_{500,001}$ is sharky while $O_{500,000}$, containing just one fewer second of intense pain, is normal. Later, we consider the possibility that the boundary between the normal and the sharky is vague.

Let us assume that O_1 is the status quo to which the other outcomes in S are to be compared. Thus, we imagine that the agent who is trying to evaluate these outcomes can grasp what O_1 is like but cannot grasp what the other outcomes in S are like. ERV entails the following two claims:

- (1) The agent cannot evaluate $O_{500,000}$ and compare it to O_1 because she cannot grasp what $O_{500,000}$ is like.
- (2) The agent can evaluate $O_{500,001}$ as worse than O_1 even though she cannot grasp what $O_{500,001}$ is like.

The conjunction of (1) and (2) entails that there are sharp discontinuities in the evaluability of outcomes. We find this implausible. In particular, we find it implausible that the agent's ability to evaluate an outcome and compare it to O_1 vanishes due to the absence of just one fewer second of pain. We call this objection to ERV *the discontinuity objection*.⁷

⁷ For similar concerns about sharp discontinuities in value, see Chang (2002) and Temkin (2012: Chapter 5).

At first glance, the discontinuity here may not seem so implausible. For example, a small physical change can make a difference to whether a certain property, such as the property of being tall or being bald, is instantiated. But the case that we are currently considering is not like these cases. According to ERV, in evaluating $O_{500,001}$, the agent can have some *doxastic attitude* about the value of this outcome and how it compares to O_1 . For example, suppose the agent can know that $O_{500,001}$ is bad and worse than O_1 . In that case, it seems she could at least have some credence that $O_{500,000}$ is also bad and worse than O_1 . After all, she has knowledge of the badness of $O_{500,001}$ *only* because in this outcome she would experience a certain amount of intense pain, and $O_{500,000}$ has only *one fewer second* of intense pain than $O_{500,001}$. But although the agent knows that $O_{500,001}$ is bad and worse than O_1 , according to ERV she is *totally clueless* about the value of $O_{500,000}$ and how it compares to O_1 . She cannot grasp the value of $O_{500,000}$. This discontinuity in evaluability is striking.

The reason for the discontinuity here is that, according to ERV, the standard of evaluation differs depending on whether the agent is trying to evaluate $O_{500,000}$ or $O_{500,001}$. In order to evaluate $O_{500,001}$, the agent does not have to cognitively model that outcome. $O_{500,001}$ contains enough intense pain that the agent can evaluate it as bad even though she cannot grasp what it is like. However, in order to evaluate $O_{500,000}$, the agent must cognitively model it. Since the agent cannot grasp what $O_{500,000}$ is like, when she attempts to cognitively model it, she ends up in a state that is similar to the state that Mary the scientist is in when, sitting in her black and white room, she tries to imagine what it is like to see red.⁸

The discontinuity objection is not the end of ERV's problems. ERV also faces what we will call *the sensitivity objection*. Consider again Fig. 1, and outcomes O_1 , $O_{500,000}$, and $O_{500,001}$. ERV implies:

- (1) The agent cannot evaluate $O_{500,000}$ and compare it to O_1 because she cannot grasp what $O_{500,000}$ is like.
- (3) The agent can evaluate $O_{500,001}$ as worse than $O_{500,000}$ even though she cannot grasp what either of these outcomes is like.

We find the conjunction of (1) and (3) very implausible. The difference in the amount of intense pain between $O_{500,001}$ and $O_{500,000}$ is *tiny* compared to the difference in intense pain between O_1 and $O_{500,000}$. It is implausible that while there being a difference of one second of intense pain is enough for the agent to be able to evaluate $O_{500,001}$ as worse than $O_{500,000}$, a difference of 499,999 seconds of intense pain is insufficient for the agent to be able to evaluate $O_{500,000}$ as worse than O_1 . Thus, when it comes to evaluating outcomes, ERV is *very sensitive* to certain small

⁸ Richard Pettigrew (2015) defends an interpretation of transformative choice that accommodates our intuition that if the agent can evaluate $O_{500,001}$, then she can also evaluate $O_{500,000}$. According to Pettigrew, the agent's uncertainty about the subjective values of different possible outcomes can be understood as uncertainty about what her utility function in those outcomes will be. This kind of uncertainty is much easier to grapple with than the kind of radical uncertainty that ERV entails, according to which one cannot even have any credence concerning the subjective values of outcomes that one has never experienced before.

differences in intense pain but *completely insensitive* to certain much larger differences in intense pain. This is *the sensitivity objection*.

The discontinuity and sensitivity objections, although related, are importantly different. The first objects to the claim that while the agent cannot evaluate $O_{500,000}$ and compare it with O_1 , she can evaluate $O_{500,001}$ as worse than O_1 . A difference of *only one second of pain* determines whether an agent who is clueless about the phenomenology of an outcome can evaluate and compare it to the status quo—there is an *implausibly sharp cut-off* in the evaluability of outcomes. The second objects to the claim that while the agent cannot evaluate $O_{500,000}$ and compare it with O_1 , she can evaluate $O_{500,001}$ as worse than $O_{500,000}$, even though the difference in pain between the first two outcomes is *much greater* than the difference in pain between the second two outcomes—the agent's evaluation of the outcomes is *not proportionally sensitive* to the variation in the amount of pain that they contain.

One may wonder whether the source of ERV's trouble is due to a failure to account for vagueness. It might be that the discontinuity and sensitivity objections arise only because we have falsely assumed a precise cut-off between the range of normal outcomes and the range of sharky outcomes. Let's suppose instead that the boundary between these two ranges is not precise but vague. Under this version of ERV, while some outcomes clearly belong to the range of normal outcomes, some other outcomes are neither determinately normal nor determinately sharky. Is recognizing vagueness sufficient to answer the discontinuity and sensitivity objections? To answer this question, we need to look further at the details.

Figure 2 below represents S under a version of ERV that accounts for vagueness. Figure 2 is similar to Fig. 1. The main difference is the presence, in Fig. 2, of the rectangular grey zone that covers $O_{500,001}$, $O_{500,002}$, and $O_{500,003}$. This zone represents the range in which the status of outcomes is vague or *borderline*—i.e. neither determinately sharky nor determinately normal. For now, we assume that there is a sharp demarcation between the zone of vagueness, the range of normal outcomes, and the range of sharky outcomes. Thus, we assume that $O_{500,000}$ is a normal outcome situated just above the zone of vagueness, and $O_{500,004}$ is a sharky outcome situated just below that zone. (We question this assumption later in our discussion).

Now, there are different ways of understanding vagueness. In the remainder of this section, we will consider the implications of ERV under three such understandings: Supervaluationism, Epistemicism, and ontic vagueness. We will argue that none of these understandings of vagueness, when combined with ERV, alleviates the discontinuity and sensitivity problems. Let's see this.

According to Supervaluationism, a sentence attributing some value to an outcome is vague if and only if it comes out true on some precisifications of the relevant evaluative predicates, but false on others.⁹ Applied to our spectrum S (Fig. 2) Supervaluationism implies that on any admissible precisifications of 'normal' and 'sharky,' O_1 and $O_{500,000}$ are normal outcomes, and $O_{500,004}$ is sharky.

⁹ For other discussions of the supervaluationist treatment of vagueness in value see Rabinowicz (2009) and Broome (2004).

Fig. 2 ERV with Vagueness

Table 1 Precisifications of ‘normal’ and ‘sharky’

	Normal	Sharky
P1	$O_1 - O_{500,003}$	$O_{500,004}$
P2	$O_1 - O_{500,002}$	$O_{500,003}, O_{500,004}$
P3	$O_1 - O_{500,002}, O_{500,003}$	$O_{500,002}$
P4	$O_1 - O_{500,001}$	$O_{500,002} - O_{500,004}$
P5	$O_1 - O_{500,000}, O_{500,002}, O_{1000,002}$	$O_{500,001}, O_{500,004}$
P6	$O_1 - O_{500,000}, O_{500,002}$	$O_{500,001}, O_{500,003}, O_{500,004}$
P7	$O_1 - O_{500,000}, O_{500,003}$	$O_{500,001}, O_{500,002}, O_{500,004}$
P8	$O_1 - O_{500,000}$	$O_{500,001} - O_{500,004}$

But because of the imprecision in our language, this is not true of $O_{500,001}$, $O_{500,002}$, and $O_{500,003}$. For each of these outcomes, on some ways of making our language more precise, it will be normal; on others, it will be sharky. Table 1 above illustrates the different possible precisifications of ‘normal’ and ‘sharky’. For each precisification there is at least one outcome that can be evaluated as bad and worse than O_1 and at least one other outcome that contains exactly one fewer second of intense pain and *cannot be* evaluated and compared with O_1 . Notice that P₃, P₅, P₆, and P₇ (the bold rows in Table 1) contain multiple instances of this kind of discontinuity. We think that these four precisifications are inadmissible, since they involve bizarre alternating patterns in evaluability. We will therefore only focus our analysis on the other precisifications, P₁, P₂, P₄, and P₈.

Figure 3 below shows the admissible precisifications, P₁, P₂, P₄, and P₈. For each of these precisifications, there is a line marking the transition between the normal and sharky outcomes.

The fact that the precisifications of the relevant predicates result in sharp discontinuities is not surprising. This is an established consequence of

Fig. 3 Graphical representation of P1, P2, P4, and P8

Supervaluationism, which can be observed by applying this view to other cases of vague predicates such as ‘tall’. But recall that the problems that arise with sharp discontinuities in the case of the evaluative predicates ‘normal’ and ‘sharky’ concern the evaluability of outcomes. And these problems do not go away simply because one makes ‘normal’ and ‘sharky’ more precise. Thus, even after the predicates have been precisified, we still get the result that the standard by which outcomes are evaluated changes depending on the presence of 1 more second of pain. Moreover, it is worth mentioning that for predicates such as ‘tall’ there is nothing analogous to the sensitivity objection. Recall that this objection arises because ERV entails that the agent’s ability to make evaluative comparisons is *very sensitive* to certain small differences in pain but *completely insensitive* to certain much larger differences in pain. On ERV, evaluative comparisons are not proportionally sensitive to differences in amounts of pain. But the relation “taller than” is always proportionally sensitive to differences in height. Moreover, comparisons of the heights of objects (tall and short) can be made no matter how those heights differ. There are no cases of the following kind: Anna is 1,50 m tall, Peter is 1,51 m tall, and Maria is 2m tall, and one can compare the heights of Peter and Anna and determine that Peter is taller than Anna but one cannot compare the heights of Peter and Maria and determine that Maria is taller than Peter. This suggests that the sensitivity objection should be understood as an objection specifically to ERV, and not to Supervaluationism *per se*.

Like Supervaluationism, Epistemicism cannot help ERV avoid the discontinuity and sensitivity objections. Epistemicism construes vagueness not as linguistic indecision but as a kind of ineliminable uncertainty. Assuming Epistemicism, there is some number n of seconds of pain such that outcome O_n in our spectrum is sharky and outcome O_{n-1} is normal, although it is impossible to know the value of n . The discontinuity and sensitivity objections thus remain unanswered. After all, these objections arise *because* the range of sharky outcomes and the range of normal outcomes are separated by the mere presence or absence of just one second of pain. Epistemicism just confirms that this is the case and adds that we cannot know exactly where the separation occurs.

Paul might claim that vagueness is not a matter of imprecision in language, or of ineliminable uncertainty, but is rather “out there” in the world. This would be a

substantial theoretical commitment on Paul's part. Very few contemporary philosophers defend this view of vagueness.¹⁰ Moreover, it seems to us that this view is vulnerable to the discontinuity and sensitivity objections. For example, ERV with ontic vagueness would have the following result. Although $O_{500,004}$ (a sharky outcome) can be evaluated as bad and worse than $O_{500,000}$ (a normal outcome), it is indeterminate whether $O_{500,003}$ (which has only one fewer second of intense pain than the sharky outcome $O_{500,004}$) can be evaluated as bad and worse than O_1 . But this is implausible, given that the difference in pain between O_1 and $O_{500,003}$ is much greater than the difference in pain between $O_{500,000}$ and $O_{500,004}$. The evaluations here are not proportionally sensitive to differences in pain.

Moreover, ERV with ontic vagueness doesn't specify what the agent should do when at least one of the outcomes that she is evaluating is borderline. Suppose that a standard is introduced in order to guide the agent in her evaluation of borderline outcomes. For example, suppose the relevant standard says that when an agent is considering two different outcomes and at least one of them is borderline, she is rationally required to cognitively model both outcomes, just as if she were evaluating and comparing two normal outcomes. Then there may be an implausibly sharp demarcation between outcomes to which this standard applies and those to which it does not apply. For example, suppose that as we have been assuming, in our spectrum $O_{500,000}$ is normal, $O_{500,004}$ sharky, and $O_{500,001}$ — $O_{500,003}$ are borderline. Then, on the proposal that we are now considering, while an agent who is deciding between O_1 and $O_{500,003}$ must cognitively model these outcomes, an exercise that leaves her completely clueless as to how they compare, that same agent can just know, without any cognitive modelling, that $O_{500,004}$ is worse than O_1 , and, moreover, that $O_{500,004}$ is worse than $O_{500,000}$. But this is basically what we found objectionable about ERV *without vagueness*—an implausibly sharp evaluative discontinuity and an implausibly volatile sensitivity to natural differences between outcomes.

One might wonder whether these problems can be avoided by appealing to *higher-order vagueness*. One might claim that, contrary to what we have assumed, it is indeterminate whether $O_{500,000}$ is borderline and indeterminate whether $O_{500,004}$ is borderline. One might also claim, for example, that it is indeterminate whether it is indeterminate whether $O_{499,999}$ is borderline, and indeterminate whether it is indeterminate whether $O_{500,005}$ is borderline, etc., etc.

Even if such claims are true, we doubt that they will be of much help. Let's say that an outcome is *second-order borderline* if it is indeterminate whether it is borderline. Now suppose that the agent is considering a second-order borderline outcome. (Assume that in this case there is no vagueness of any order higher than 2). How should the agent evaluate this outcome? Here is one proposal. The relevant standard of evaluation for *borderline* outcomes requires that the agent cognitively model them, so for any *second-order borderline* outcome O , the relevant standard requires that the agent make a *second-order evaluation* by cognitively modelling an

¹⁰ Notable exceptions include Van Inwagen (1990) and Tye (1990). For a more recent interpretation of ontic vagueness, see Barnes (2010).

outcome in which she cognitively models O. Extrapolating to higher orders, we can say that for any n -order borderline outcome, where $n > 1$, the relevant standard requires the agent to make an n -order evaluation by cognitively modelling an outcome in which she uses cognitive modelling to make an $(n-1)$ -order evaluation.

The problem with this proposal is that using cognitive modelling for higher-order evaluations is a redundant epicycle. We are interested in cases of evaluation in which the agent is clueless about what the outcomes are like phenomenologically. But if an agent is clueless about what a certain outcome is like, then presumably she is also clueless about what it is like to cognitively model that outcome. After all, cognitively modelling an outcome involves estimating what it is like. Thus, with respect to outcomes such that the agent is clueless about what they are like, using cognitive modelling for higher-order evaluations would only exacerbate her cluelessness.¹¹

We think that the problem just considered points toward an alternative proposal: in cases of n -order vagueness, where $n > 1$, rather than go through some number of redundant epicycles, the agent should just act exactly as she would in any case of first-order vagueness. In other words, the standard of evaluation for cases of higher-order vagueness is just the standard of evaluation for cases of first-order vagueness. Assuming that in cases of first-order vagueness cognitive modelling is required for the agent's first-order evaluation, this is also true in cases of higher-order vagueness; in such cases, no higher-order evaluation is required.

But this proposal faces problems of its own. Suppose that in our spectrum S there is a sharp demarcation between outcomes that are n -order borderline for some $n \geq 1$ and outcomes that are not n -order borderline for any n . Then, on the current proposal, in S there will be a sharp demarcation between outcomes to which the standard requiring cognitive modelling applies and outcomes to which it doesn't. But this is precisely what gets ERV into trouble in the first place.

Suppose instead that *every outcome* in S is n -order borderline for some $n \geq 1$. Then on the current proposal, for any outcome in S, in order to evaluate that outcome, the agent must cognitively model it. But this is incompatible with one of ERV's assumptions, namely that at least some outcomes in S, for example those involving hours and hours of extreme pain are sharky—i.e. are such that one doesn't *need* to cognitively model them but can just know (or grasp) that they are bad no matter what they are like.

At this point, one might wonder whether the problems that we have discussed can be avoided by abandoning cognitive modelling as the standard of evaluation for

¹¹ A further problem with the proposal may arise depending on how one understands higher-order vagueness. For example, suppose one adopts Timothy Williamson's view on which any sentence of the object language has higher-order vagueness only if it has n -order vagueness for every natural number n . (See Williamson 1999, especially pp. 136–138). Then with respect to sentences with the evaluative predicates 'normal', 'sharky', and 'borderline', any case of higher-order vagueness will be a case in which there is vagueness *all the way up*. On the proposal we are now considering, this means that in any case of higher-order vagueness the agent has no way of getting started with her evaluations. For in order to make her second-order evaluation, she must first make a third-order evaluation; and in order to make this third-order evaluation, she must first make a fourth-order evaluation, and so on *ad infinitum*. However, Williamson's view of higher-order vagueness assumes classical logic. Thus, while the problem of an infinite regress may arise on theories of vagueness that assume classical logic, such as Supervaluationism and Epistemicism, it may not arise on theories that reject classical logic.

borderline outcomes and adopting some alternative standard. Although we cannot rule out this possibility, the onus is on those who accept both the Shark Claim and the Prior Experience Claim to demonstrate it. However, we *suspect* that any reasonable alternative to a standard that requires cognitive modelling will face either the discontinuity and sensitivity problems or problems that are structurally analogous to them. In Sect. 4, we show that Paul's proposed alternative standard for cases of transformative choice faces such problems, and we cannot think of any other reasonable standard that avoids them.

To sum up, the Experiential Range View (ERV), which restricts the application of the Prior Experience Claim in order to accommodate the Shark Claim, involves an implausible discontinuity in the evaluability of outcomes. Moreover, on this view, evaluative comparisons of outcomes are not proportionally sensitive to variation in the underlying features of these outcomes. Finally, these problems don't seem to go away on the assumption that the boundary between the range of normal outcomes and the range of sharky outcomes is vague.

3 The richness assumption

In this section we defend a crucial assumption on which our arguments in the previous section depend, namely that the number of types of experience is sufficiently large to fill in S. We call this The Richness Assumption. Those who think that the variety of types of experience is *too limited* to fill in S reject The Richness Assumption.

We think that one should not underestimate the enormous diversity of types of human experience, both throughout history and at any given time. Paul mentions a number of categories that determine the subjective experience of someone's life, some of which include gender, race, or affluence.¹² The following is a non-exhaustive list of categories that expands on Paul's idea: gender, ethnicity, profession, parenthood status, personality, religion, mental health status, disability status, sexual orientation, and addiction status. Each of these categories could also be subdivided into a non-exhaustive list of more specific sub-categories, where the sub-categories within any given category determine very different subjective experiences. For example, consider the following subcategorization:

1. *Gender* man, woman, trans-man, trans-woman, non-binary
2. *Ethnicity* African-American, African, Northern-European, Latin-American, Asian-American, Southern-European
3. *Profession* lawyer, violinist, teacher, CIA operative, politician, doctor, pilot, athlete, actor/actress, writer, philosopher, farmer, special needs teacher, bank CEO, astronaut, gravedigger, brain surgeon, priest, professional climber, sex worker
4. *Parenthood* Parent, non-parent
5. *Personality* Introvert, extrovert

¹² For more on this, see Paul (2014: 8).

6. *Religion* non-religious, Islamic, Christian, Jewish, Buddhist, Hindu, Tribal religion
7. *Mental health* depressed, non-depressed, psychopathic-depressed, psychopathic-non-depressed
8. *Disability* physical, intellectual, emotional, physical & intellectual, physical & emotional, physical & blind, physical & deaf, non-disabled
9. *Sexual Orientation* gay, straight, bi-sexual
10. *Addiction* addicted, non-addicted

The combination of these sub-categories corresponds to a huge number (3,225,600) of possible life-experiences (or outcomes) that, we think, differ greatly from each other with respect to their phenomenological content. For example, consider the following pair of lives, L1 and L2:

L1 The life of a woman who is also African-American, a lawyer, physically disabled, and religious.

L2 The life of a man who is also Northern-European, a lawyer, non-physically disabled, and non-religious.

It seems obvious that L1 and L2 are phenomenologically very different. For example, other things being equal, the person who leads L1 might experience more disadvantage, and be more of a target of discrimination than the person who leads L2. And for someone who has lived only L1 (or L2), the experience of living L2 (or L1) would be epistemically inaccessible, given that the precise phenomenology of the experiences associated with these lives is very different.

Even lives that differ only in one of the sub-categories on our list could be phenomenologically different enough that a person living one of these lives could not grasp what it is like to live any of the others. For example, consider:

L3 The life of a woman who is also Southern-European, a lawyer, religious, and **nondisabled**.

L4 The life of a woman who is also Southern-European, a lawyer, religious, and **disabled**.

We think that even though L3 and L4 differ only in one sub-category, this difference alone could make the two lives experientially different enough that someone living L3 could not grasp what it was like to live L4 (and vice versa).

The Richness Assumption can be supported even further. The experiences that can be constructed out of the categories listed above represent only a small fraction of the total number of humanly possible experiential types. Lives in the (more or less distant) future may turn out to be experientially unimaginable for us. Technological advances are the main reason for this gap in experiential knowledge. People in 400 B.C. probably could not imagine what it is like to fly by plane or use a computer. Further technological advances will make possible some things that are currently unimaginable. For example, the acquisition of new sensory abilities would enable humans to access a whole set of new types of experiences that would otherwise have remained unknown to us.

Finally, we think that humanly possible experiential types represent only a small fraction of the total number of possible experiential types. For example, Paul considers the lives of imaginary creatures such as vampires (Paul 2014). Once we consider not only the kinds of human experiences mentioned above but all metaphysically possible experiences, the Richness Assumption seems extremely compelling.

To sum up, we find it plausible that human experience is diverse enough to satisfy the Richness Assumption, and even more plausible that experience in general (including non-human experience) is diverse enough to satisfy this assumption.

4 The shark problem and the revelation approach

We have so far focused on the fact that restricting the application of the Prior Experience Claim to accommodate the Shark Claim leads to implausible discontinuities in the evaluability of outcomes, and that this constitutes a strong reason to reject one of these two claims. We have referred to this as the shark problem.

In this section, we consider whether the Revelation Approach proposed by Paul helps avoid the shark problem. According to this approach, when the possible outcomes of one's decision contain experiences of a kind that one has never had before, one can make a rational choice based on whether one wants to have a revelation in which one discovers what it is like to have experiences of that kind. This approach introduces a new concept, that of *revelatory value*. When we apply this approach, the agent does not attempt to cognitively model the outcomes that she must evaluate. The agent can evaluate, compare, and ultimately choose an outcome based on its revelatory value alone.

Indeed, Paul argues that in these cases one's decision is rational only if it is based on one's desire to find out (or not find out) what it would be like for her to have these experiences:

[T]he proposed solution is that, if you are to meet the normative rational standard in cases of transformative choice, you must choose to have or to avoid transformative experience based largely on revelation: you decide whether you want to discover how your life will unfold given the new type of experience. ... If you choose to undergo a transformative experience and its outcomes, you choose the experience *for the sake of discovery itself*, even if this entails a future that involves stress, suffering, or pain (Paul 2014: 120, emphasis added).

One may argue that the Revelation Approach can also be applied to choices between non-sharky outcomes in the spectrum S.¹³ The approach might say the

¹³ Here we use the term 'non-sharky outcomes' rather than 'normal outcomes'. According to our definition of 'normal outcomes', one cannot make a rational choice between such outcomes unless one has had experiences of the kind that occur in them. But on the Revelation Approach, one can make a rational choice between certain outcomes even if one has not had experiences of the kind that occur in them. Therefore, such outcomes are not normal outcomes on our definition.

following. Each non-sharky outcome in S has a revelatory value, and one can choose between the non-sharky outcomes in S based solely on their respective revelatory values. However, when comparing a sharky outcome with a non-sharky outcome one can just know that the former is bad and worse than the latter, without having to consider their respective revelatory values.

One might think that by introducing revelatory values, one can avoid the discontinuity and sensitivity objections. The solution would be to claim that all outcomes in S can be directly evaluated, either on the basis of their revelatory value or on the basis of one's knowledge of their badness (as in the case of sharky outcomes).

Unfortunately, this approach faces versions of the discontinuity and sensitivity objections. Assume for now that there are no borderline outcomes in S , i.e. that for every outcome in S it is either determinately true that this outcome is sharky, or determinately false that it is sharky. Then there will be a sharp discontinuity between outcomes to which one must apply the Revelation Approach and outcomes to which one need not apply this approach. Take for example $O_{500,000}$, which is not sharky, and $O_{500,001}$, which is sharky. On the current proposal, one must evaluate $O_{500,000}$ solely on the basis of its revelatory value, but one can evaluate $O_{500,001}$ directly, without having to consider whether one wants to have a revelation in which one finds out what it is like. This is bizarre. Suppose you decide that the only way to evaluate $O_{500,000}$ is to ask yourself, "Am I the kind of person who would like to learn what $O_{500,000}$ is like?". Next, suppose you are considering outcome $O_{500,001}$, which contains *only* one more second of pain than $O_{500,000}$. On the current proposal, your evaluative task is now straightforward: you need not consider whether you would like to have a revelation where you learn what $O_{500,001}$ is like. You know that this outcome is, as Paul would say, "bad whatever it is like", regardless of its revelatory value. But if $O_{500,000}$ can be evaluated solely on the basis of its revelatory value, why would the addition of one more second of pain make revelatory value irrelevant?

To see that the approach is vulnerable to a version of the sensitivity objection, consider the following. Neither O_1 nor $O_{500,000}$ is sharky. So, on the current proposal, when comparing O_1 and $O_{500,000}$, which differ greatly in their respective amounts of pain, one must apply the Revelation Approach. On the other hand, when comparing $O_{500,000}$ with $O_{500,001}$, a sharky outcome containing one more second of pain, one need not consider revelatory value; instead, one *just knows* that $O_{500,001}$ is bad and worse than $O_{500,000}$. The criterion for applying the Revelation Approach is extremely sensitive to small differences in pain, but strangely insensitive to larger differences in pain.

Finally, for reasons that we saw in Sect. 2, we do not think that appealing to vagueness will help with these problems. Suppose that some outcomes in S are borderline—it is neither determinately true that they are sharky nor determinately false that they are sharky—and that there is no higher-order vagueness. Should the agent follow the revelation approach when evaluating borderline outcomes, but not

when evaluating sharky outcomes? If so, then in *S* there will be a sharp demarcation between outcomes to which the Revelation Approach applies and outcomes to which it doesn't.

Otherwise, if there is higher-order vagueness, we need to consider how the agent should evaluate higher-order borderline outcomes. Is the agent rationally required to decide whether to have a *higher-order revelation*, i.e. a revelation in which she finds out what it is like to find out what some outcome is like? But this standard seems needlessly complicated. It is also practically impossible to follow for *n*-order revelations where *n* is very large.

On the other hand, perhaps the agent is required to evaluate higher-order borderline outcomes exactly as she does borderline outcomes—i.e. by deciding whether to have a first-order revelation in which she finds out what one of those outcomes is like. But if there is a sharp demarcation in *S* between outcomes that are *n*-order borderline for some $n \geq 1$ and outcomes that are not *n*-order borderline for any *n*, then this will just lead to the same old problems. And if every outcome in *S* is *n*-order borderline for some $n \geq 1$, then we get the absurd result that the Revelation Approach applies even when the agent is considering outcomes that, intuitively, are sharky. Presumably, the agent can know that such outcomes are bad; she doesn't need to consider whether to have a revelation in which she finds out what they are like.

5 Conclusion

We have argued that restricting the Prior Experience Claim in order to accommodate the Shark Claim leads to implausible discontinuities in the evaluability of outcomes. This is the shark problem. We think that in the face of the shark problem, we should reject either the Prior Experience Claim or the Shark Claim. Our view is that the Shark Claim is significantly more plausible than the Prior Experience Claim. We all know that being eaten alive by sharks is bad even if we don't know what that is like.

Acknowledgements We are grateful to the participants in the PPE seminar of the Institute for Futures Studies (IFFS), Stockholm, April 2017; the 2017 Stockholm June Workshop in Philosophy, Stockholm University; and the 2018 conference *New Perspectives on Transformative Experience*, University of Gothenburg, for their helpful comments on this paper. Special thanks L.A. Paul, Vuko Andric, Paul Bowman, Krister Bykvist, Erik Carlson, and Luke Elson for helpful comments on earlier drafts of this paper. In addition, we would like to thank Nomy Arpaly, Gustaf Arrhenius, Johan Brännmark, Lisa Furberg, Anna-Sofia Maurin, Alain Pe-Curto, H. Orri Stefansson, and Irina Vartanova for her very helpful assistance with the diagrams in this paper.

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

permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Funding Funding was provided by Riksbankens Jubileumsfond (SE) (Grant No. M170372:1).

References

- Barnes, E. (2010). Ontic vagueness: A guide for the perplexed. *Nous*, 44(4), 601–627.
- Broome, J. (2004). *Weighing lives*. Oxford: Oxford University Press.
- Bykvist, K., & Stefansson, O. (2017). Epistemic transformation and rational choice. *Economics and Philosophy*, 33(1), 125–138.
- Chang, R. (2002). The possibility of parity. *Ethics*, 112, 659–688.
- Paul, L. A. (2014). *Transformative experience*. Oxford: Oxford University Press.
- Paul, L. A. (2015). Transformative experience: Replies to Pettigrew, Barnes and Campbell. *Philosophy and Phenomenological Research*, 91(3), 794–813.
- Pettigrew, R. (2015). Transformative experience and decision theory. *Philosophy and Phenomenological Research*, 91(3), 766–774.
- Rabinowicz, W. (2009). Incommensurability and vagueness. *Proceedings of The Aristotelian Society*, 83, 71–94.
- Sullivan, M. (2018). *Time biases: A theory of rational planning and personal persistence*. New York: Oxford University Press.
- Talbott, W. (2016). Transformative experience. *Analysis*, 76(3), 380–388.
- Temkin, L. (2012). *Rethinking the good: Moral ideals and the nature of practical reasoning*. New York: Oxford University Press.
- Tye, M. (1990). Vague objects. *Mind*, 99, 535–557.
- Van Inwagen, P. (1990). *Material beings*. Ithaca: Cornell University Press.
- Williamson, T. (1999). On the Structure of Higher-Order Vagueness. *Mind*, 108, 127–143.

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