Evaluating Transformative Decisions

Vladimir Chituc, L.A. Paul, M.J. Crockett

Abstract

Recent philosophical work has taken interest in the decision-theoretic problems posed by transformative experiences, or experiences that are epistemically revelatory and life-changing (like becoming a parent). The problem is roughly as follows: if we cannot know what it's like to be a parent (its subjective value) before actually becoming one, then how are we to decide whether to become one? This topic has received recent empirical attention, some of which has challenged the central importance of subjective value for transformative decision-making. Here, we present empirical work that suggest these findings can be explained by the *evaluability bias*, in which people weigh decision criteria not based on their importance, but how easy they are to evaluate. Participants not only find subjective value important, but they report willing to pay a great deal of money to get this information. Furthermore, participants who were most uncertain about whether to undergo a transformative decision we most likely to report interest in seeking out information about subjective value. We conclude by considering the philosophical and empirical implications of this work.

Keywords: Transformative Experience; Experimental Philosophy; Judgment and Decision Making; Subjective Value

Introduction

How do we make decisions about novel experiences? Oftentimes, we are able to combine and simulate relevant past experiences as an approximation, just as we might get a decent sense of how an unfamiliar dish will taste based on the ingredients listed on a menu (see e.g.: Barron, Dolan and Behrens, 2013; McCoy et al., 2019; Kappes and Morewedge 2016; Williamson 2016). This straightforward picture breaks down, however, in the case of transformative experiences.

Transformative experiences—like the gain or loss of sensory capacities, becoming a parent, emigrating to a country with a culture very different from one's own, or fighting in a war—are epistemically revelatory and life-changing. By undergoing a transformative experience, we learn what a new kind of experience is like, and this gives us new abilities to imaginatively represent and accurately simulate possible states of affairs involving this experience (Jackson 1986, Lewis 1990, Paul 2014, Paul 2015a, Paul 2015b). Such discoveries can then lead the person to change in some deep and personally fundamental way; for example, some of their core personal preferences or values might change. Thus, transformative experiences are defined as experiences that are both *epistemically* and *personally* transformative.

Transformative experiences pose a problem for the straightforward solution to novel decisions we described above. Paul (2014) argues that this is because such decisions hinge on *subjective value,* leading to a decision-theoretic bind. Here, *subjective values* are experientially grounded values that include an assessment of the nature of what it's like to live "in this," as John Campbell (2015) puts it.[1] The difficulty, however, is that a person must decide whether to undergo the transformative experience without knowing the subjective value of that experience, since they have not had the epistemic and personal changes that the experience entails.

For example, to know the subjective value of being a loving parent, a person has to actually be a loving parent. To be a loving parent, however, profoundly changes the way someone experiences, and thus evaluates, their life in many other respects, including how they understand themselves in relation to other people, how they evaluate career and life goals, and the way they experience vulnerability and loss, and so on. This is an important reason why, in real-world contexts, testimony from other parents about how one's life will change is famously insufficient for prospective parents to be able to grasp what their new, post-partum, lives will be like.

In recent empirical work on transformative experiences, Reuter and Messerli (2018) argue that the problem posed by transformative experiences is not actually a problem at all, arguing that people do not care about subjective value when making transformative decisions. If this is the case, and if subjective values are not important in real-life decision making, then the inaccessibility of such values is not a threat for practical decisions in transformative contexts.

In this work, we contribute to the rapidly expanding body of experimental research on transformative decision making, and in the process, we mark out a novel empirical interpretation for assessments of subjective value in transformative contexts. Our argument proceeds as follows. We start by providing an overview and critique of Reuter and Messerli (2018)'s claims about subjective value, arguing that, often, people attach more weight to decision

[1] Note: the use of *subjective value* here and in the broader literature on transformative decisions is distinct from how the term is used by economists and neuroscientists (c.f.: Kable and Glimcher, 2007).

criteria that they can know — a phenomenon known as the "evaluability bias". Therefore, if people cannot know the subjective value when making a decision, they are unlikely to attach subjective value much weight, even if they care about it. Next, we describe an empirical study that supports this interpretation of Reuter and Messerli's findings. Our results suggest that evaluability bias is indeed at play in transformative decisions, and we find support for the hypothesis that subjective value matters to people when they make transformative decisions.

The Evaluability of Subjective Value

Reuter and Messerli (2018) develop a choice model for making rational transformative decisions that specifies "weights" on decision criteria. In the case of deciding whether to have a child, for example, these criteria include the preferences of one's partner, the costs of having a child, and the subjective value of having a child. Accordingly, the overall utility of the decision to have a child is determined by multiplying the extent to which each criterion favors having a child by the weight placed on that criterion, and summing across these weighted values. Their model indicates that as long as the weight on subjective value ($w_{\text{whatitslike}}$) is less than the sum of the weights on all other decision criteria (w_{partner}, w_{costs}, and so on) then transformative decisions can be rational. That is, if $w_{\text{whatitslike}} < 0.5$, then subjective value is not "the central" decision criterion and can be safely cast aside in the decision, if, as they claim, it is "not important enough to influence the decision process" (p. 11).

To get to this claim, they conducted empirical studies in attempts to discover the actual weights people attach to various decision criteria, including subjective value, when making transformative decisions. Participants were asked to consider three different types of transformative experiences—having a child, becoming a vampire, and living on Earth under alien rule—and to indicate how much they would weigh each of six different decision criteria in order to arrive at a decision. Participants did not weigh subjective value higher than they weighted other decision criteria, and based on these findings, Reuter and Messerli conclude that "...the importance of the subjective value is often much lower than has so far been assumed by Laurie Paul [sic] and other scholars working on this topic" (p. 24).

We disagree with this conclusion and suggest there is a more interesting phenomenon in play. Taking Reuter and Messerli's data at face value, if it is indeed the case that people report they do not weigh subjective value more strongly than other criteria in transformative decisions, does it follow that they think subjective value is not an *important* criterion in those decisions?

We think not. Participants in these studies were asked "How would you arrive at a decision?", which is fundamentally a question about how people think they would weigh different criteria in their decision process, and research on decision-making shows that people weigh decision criteria in proportion to how easy they are to evaluate (Bazerman, Loewenstein, & White, 1992; Hsee, 1996; Hsee & Zhang, 2004, 2010; Caviola et al., 2014). This phenomenon, known as the "evaluability bias," shows that people place a low weight on certain decision criteria because they are difficult to evaluate, not because they consider such criteria unimportant.

Consider two diamond shoppers: an expert gem trader and a naive shopper who knows very little about gemstones. The expert is likely to heavily weigh several criteria when deciding which stone to buy and how much to pay for it, including cut, color, clarity, and carat (size). The more knowledge the expert has about these criteria (e.g., how to evaluate the clarity of a stone), the more sensitive she will be to variations in these criteria when determining how much a stone is worth to her (e.g., decreases in clarity among stones will track more closely with decreases in value).

The naive shopper, in contrast, lacks the ability to evaluate some of these criteria. As a consequence, he will be insensitive to variations among stones on those criteria he is unable to evaluate, and base his decision only on those criteria that are easy to evaluate (e.g., carat size). If he does not know how to evaluate clarity, for instance, he will not demand a lower price for a stone with poor clarity. But that doesn't mean he does not care about clarity and would not value clarity if he were taught how to evaluate it; it just means that he lacks the knowledge necessary to evaluate clarity, and therefore cannot weigh clarity in his decision process. When he decides which diamond to buy, he will not assign a significant role to judgments of clarity. He may even make a suboptimal decision: an unscrupulous seller, detecting the naive shopper's lack of knowledge, could charge him a higher price for a large stone with poor clarity than he could charge the expert gem trader.

The naive shopper's over-reliance on carat size in his decision process is an instance of the evaluability bias. Classic work by Kahneman and colleagues demonstrated that what people predict will make them happy (predicted utility) and the values people base their decisions on (decision utility) often substantially differ from what actually makes them happy (experienced utility; Kahneman, 2000; Kahneman & Snell, 1990, 1992; Kahneman, Wakker, & Sarin, 1997; c.f. Hsee & Zhang, 2004). That is, there is extensive empirical evidence showing that when people are asked to evaluate multiple decision criteria and predict which ones will make them happy in the future, they are notoriously bad at doing this. Hsee & Zhang (2010) comprehensively review studies showing that when deciding between two options that differ on criteria that are easy and difficult to evaluate, decision-makers inappropriately over-weight criteria that are easy to evaluate. For example, when deciding between an interesting job that pays $60,000/year and a tedious job that pays $70,000/year, people overestimate the impact that the $10k difference in salary will have on their future happiness, because salary differences are easy to evaluate (whereas the distinction between an "interesting" and a

"tedious" job is less easy to evaluate). This can lead to suboptimal choices, such as choosing the tedious $70,000 job over the interesting $60,000 job, even if the latter would bring them more happiness (Hsee & Zhang, 2004).

Thus, subjective value may well be important, but it may not be a criterion that people are easily able to evaluate when contemplating a transformative decision. Thus, they may not weigh it heavily when they decide. Consider the fictional vampire case from Paul (2014): imagine that you have a one-time only chance to become a vampire.[2] By definition, as a mere human, you know that you can't understand what it's like to be a vampire until you actually become one. If you care very much about what it would be like to be a vampire and think it should play a major role in your decision, how are you to make your choice? Contemporary research on decision making and the evaluability bias suggests that, when making your decision, you are likely to place a low weight on the subjective value of life as a vampire in your decision process—not because you don't care about what it's like, but simply because you know you can't know this ahead of time. That is, you know you can't properly consider the subjective value in your decision-making process, so you don't. Instead, given these practical constraints, you focus on what you *can* know. Perhaps, as you decide, you accord significant weight to criteria such as whether friends and relatives have also become vampires, or to other factors that you can assess, such as the value of sunning yourself on a beach or looking fabulous in black. Despite the fact that the subjective value of being a vampire is very important to you, you don't weigh it heavily when contemplating your transformative decision. The relevance of this point for Reuter and Messerli's (2018) work should be clear: even if, when considering whether to become a vampire, people *do not* weigh subjective value higher than they weigh other decision criteria, this does not mean they *would not* consider subjective value to be an important or even a central criterion.

In sum, participants in Reuter and Messerli's (2018) experiments may suffer from the evaluability bias. They may report that subjective value is not central in how they'd make a transformative decision, but this is not necessarily because subjective value is not an important criterion; rather, it may be because subjective value is difficult to evaluate. Reuter & Messerli's interpretation of their results fails to take into account the well-confirmed possibility of evaluability bias.

Our line of reasoning leads to a straightforward prediction: if subjective value is not central in the decision process because it is difficult to evaluate (as opposed to being unimportant), then people should value the opportunity to gain information about it. To test our prediction, we ran a replication of Reuter & Messerli (2018), with some additional, key questions. We describe this work in the following section.

[2] Modern vampires drink artificial blood, so assume any major ethical concerns can be set aside.

Procedure and Results

We recruited 100 adults (mean age: 27; 51 females, 49 males) from the crowdsourcing website Prolific Academic. Because our survey concerned the decision of whether to have a child, we recruited only participants aged 18-40 who did not already have children. This sampling procedure ensured that our dataset only included participants for whom the decision to have a child was (a) transformative (as it may not be for those who already have children), and (b) likely to be one they were actively contemplating (as people outside this age range are less likely to be considering having a child). The procedure was approved by the Yale University Institutional Review Board, and participants were paid at a rate of $7.65/hour for completing the survey. After providing informed consent, participants were presented with the same parenthood scenario and questions as in Reuter and Messerli (full survey text is available online in Appendix 1 at https://osf.io/dbc8p/):

> *Imagine considering becoming a parent and having to decide whether or not to have a child. How would you arrive at a decision?*

Participants rated six decision criteria based on how important they were on a scale from 0 (not at all important) to 10 (extremely important). These decision criteria included discussing the decision with one's partner ("partner"), financial costs ("cost"), whether becoming a parent is consistent with life goals ("consistent"), what it will feel like to have the experiences and emotions of being a parent ("subjective value"), whether they are happy to undergo changes to personality ("openness"), and reading about the pros and cons of having a baby ("reading"). See Table 1 for participants' average ratings of these criteria. We replicated the findings of Reuter and Messerli in terms of the rank ordering of the decision criteria, suggesting that our participant samples are equivalent and the conclusions we draw from our dataset are likely to apply to Reuter and Messerli's dataset, as well.

We also added a number of key questions to specifically test our hypothesis: that subjective value doesn't weigh heavily when making transformative decisions because it is difficult to evaluate, not because it is unimportant. To test this, we presented participants with the following scenario:

> *Next, imagine that you have the opportunity to be transported into a possible future where you have your child. You would get to spend 24 hours experiencing what it is like for YOU to be a parent. When you come back from being transported, no time will have passed in the present, but you will have perfectly vivid memories of this daylong experience of your life as a parent.*

We asked participants several questions regarding this scenario. First, we asked them if they would take this opportunity. Second, we asked them to explain why or why not in their own words. Third, we asked them how much money they would be willing to pay for this opportunity. Following this, we asked a number of additional demographic questions, including whether or not they wanted to have children, and how certain they were about this preference.

Our results strongly support (i) the claim that the evaluability bias may be in effect for transformative decisions, and (ii) the claim that people value the opportunity to gain information relevant to subjective value. First, 76% of our participants indicated they would take this opportunity. Of these participants, 91% indicated they would be willing to pay for the opportunity. Willingness to pay was highly skewed (mean $137,450; median $150), with some participants indicating they would be willing to pay very high amounts (up to $10 million).

Table 1: How people weigh different criteria in parenthood decisions. The table below displays the mean rating and standard deviation for each criterion in Reuter and Messerli (2018) and our replication.

Criterion	R&M original	Replication
Partner	8.66 (2.52)	9.27 (1.56)
Costs	8.26 (2.24)	8.95 (1.55)
Consistency	7.85 (2.69)	8.7 (1.61)
Subjective value	7.68 (2.26)	8.01 (2.08)
Openness	6.43 (3.04)	7.51 (2.38)
Reading	5.45 (3.34)	6.96 (2.93)

Furthermore, willingness to take the transporter varied both as a function of whether participants indicated they wanted children and how certain they felt about this (Figure 1). Only 47% of participants who were very certain that they did not want children said they would take the transporter. In contrast, those who were most willing to take the transporter were those who reported that they did want children, but were uncertain about it (96%). Of those participants who reported a preference for having children one way or another, being uncertain about that preference was associated with a significantly higher likelihood of wanting to take the transporter ($t(67)=2.31$, $p = 0.02$), controlling for preference. That is, regardless of whether participants leaned toward wanting children or not, being uncertain about that preference significantly increased the likelihood of wanting to take the transporter. This supports our claim that people want to seek out information relevant for assessing what it's like to be a parent, particularly when they are uncertain about this transformative decision.

Second, participants' explanations for why (or why not) they would take the transporter support our additional claim that, in particular, gaining information *relevant to assessing subjective value* is valuable to them. Of those participants who reported they would take the transporter, most of them explicitly mentioned they would take it to see what it is like to be a parent, i.e., to gain information relevant to subjective value. In addition, many of the participants who reported that they would take the transporter explicitly mentioned that they thought the experience would help them make a better decision about whether to have a child. Meanwhile, those participants who reported that they would not take the transporter gave rather different explanations, for example mentioning concerns that the experience would be negative, or explicitly mentioning that they are certain they don't want kids and therefore have nothing to gain from the experience. Many of our participants reported that they sought further information about how they'd evaluate many other changes in their life that would flow from becoming a parent, including (but not limited to) changes in other relationships, career goals, and so on.

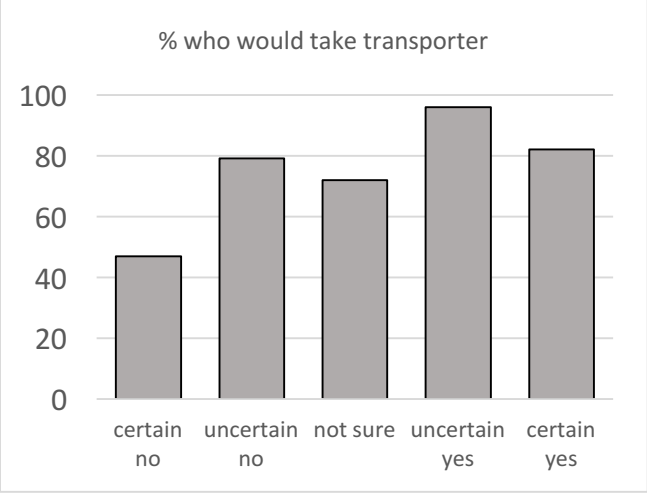

Figure 1: Percentage of participants who would take a transporter to have a daylong experience of their life as a parent, conditioned on wanting to have a child (yes, no, or not sure) and uncertainty around that preference (certain or uncertain).

Overall, these data reinforce the claim that people think that having a child leads to new experiences that can't be anticipated. They further support the claim that having these

new experiences could profoundly change the assessment of many other aspects of one's life. Raw explanation data and coding are provided online in Appendix 2 at https://osf.io/dbc8p/.

Conclusion

We've shown that the evaluability bias may be in effect for transformative decision-making. In our study, the vast majority of participants valued the opportunity to learn information relevant to subjective value. That is, they valued the opportunity to learn information about what it would be like to be a parent. Our results confirm previous empirical work on the evaluability bias while strongly supporting the possibility that people place a low weight on subjective value in transformative decisions because they lack knowledge about subjective value—and they understand, at least implicitly, that they lack this knowledge. These findings support our hypothesis that participants do not place a low weight on subjective value when they are deliberating because they do not think it is important. Instead, participants place a low weight on subjective value because it is difficult to evaluate.

At this point, a critic might argue something like the following. Perhaps our pattern of results has little to do with subjective value, since participants are receiving information about subjective value from an extremely novel and unique experience. It could be that participants just want to live a day in the life of their future selves, are willing to pay for that opportunity, and gaining knowledge about subjective value is merely incidental. We think this interpretation is unlikely. To start, a mere preference for novelty seeking would be unable to explain why those who were most uncertain about whether or not to have a child were most willing to take the transporter.

Even so, we aim to conduct follow up work to rule out this possibility and more directly support our hypotheses. First, we might ask whether participants would take the transporter for a decision where there isn't the same inability to know the subjective value of an outcome. For example, would the pattern of results look different for parents deciding to have a second child? Since these parents would already have a sense of the subjective value of parenthood (and thus could evaluate it), we would predict that they would weigh subjective value more highly. Accordingly, we would expect them to be less willing to take the transporter. If our results are driven instead by the novelty of time travel, we should expect no such pattern. Second, if the pattern of results obtained by Reuter & Messerli (2018) is explained by evaluability bias, we might expect the weights to change following the transporter manipulation. If participants had access to information gained from taking the transporter, we should expect them to weigh subjective value more heavily afterwards. If subjective value plays little role in the decision to become a parent, participants would be unlikely to incorporate information from the transporter into their decision.

Our results are important for at least three reasons. First, the role and importance of subjective value is much discussed in the philosophical literature surrounding transformative experience, for example, in Paul (2015b), Campbell (2015), Kauppinen (2015), Kind (2020), and Arpaly (2020). Our work highlights the significance of these discussions. Second, the work overturns prior theoretical interpretations of this type of data (Reuter and Messerli 2018). Third, our ideas about evaluability bias and its relation to transformative decision making develop new connections between philosophy, behavioral economics and social psychology, potentially opening up new avenues for interesting interdisciplinary research.

We hope that future empirical research will continue to explore paths between practical deliberation, formal epistemology, and social psychology, and in particular that the expanding interest in conducting empirical investigations of transformative decision making will continue to shed light on the philosophical questions about the self and value at the heart of practical deliberation and rational choice.

Acknowledgments

The authors would like to thank Vivian Fung for her research assistance, as well as Endure McTier, Charlotte Austin, and other members of the Crockett Lab for their helpful feedback.

References

Arpaly, N. (2020). What Is It like to Have a Crappy Imagination?. *Becoming Someone New: Essays on Transformative Experience, Choice, and Change*, 122.

Barron, H. C., Dolan, R. J., & Behrens, T. E. (2013). Online evaluation of novel choices by simultaneous representation of multiple memories. Nature neuroscience, 16(10), 1492.

Bazerman, M. H., Loewenstein, G. F., & White, S. B. (1992). Reversals of preference in allocation decisions judging an alternative versus choosing among alternatives. Administrative Science Quarterly, 37(2), 220–240.

Campbell, J. (2015). "L. A. Paul's Transformative Experience", Philosophy and Phenomenological Research, 91, 787–793.

Caviola, Lucius, Nadira Faulmüller, Jim AC Everett, Julian Savulescu, and Guy Kahane. "The evaluability bias in charitable giving: Saving administration costs or saving lives?." Judgment and decision making 9, no. 4 (2014): 303.

Hsee, C. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. Organizational Behavior and Human Decision Processes, 67(3), 247–257.

Hsee, Christopher K., and Jiao Zhang. "Distinction bias: Misprediction and mischoice due to joint evaluation." *Journal of personality and social psychology* 86, no. 5 (2004): 680.

Hsee, Christopher K., and Jiao Zhang. "General evaluability theory." *Perspectives on Psychological Science* 5, no. 4 (2010): 343-355.

Jackson, Frank. 1986. "What Mary Didn't Know." The Journal of Philosophy 83 (5): 291–295.

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, *10*(12), 1625–1633. https://doi.org/10.1038/nn2007

Kahneman, D. (2000). Experienced utility and objective happiness: A moment-based approach. In D. Kahneman & A. Tversky (Eds.), Choices, values and frames (pp. 673–692). Cambridge, England: Cambridge University Press.

Kahneman, D., & Snell, J. (1990). Predicting utility. In R. M. Hogarth (Ed.), Insights in decision making (pp. 295–310). Chicago: University of Chicago Press.

Kahneman, D., & Snell, J. S. (1992). Predicting a changing taste: Do people know what they will like? Journal of Behavioral Decision Making, 5, 187–200.

Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. Quarterly Journal of Economics, 112, 1937–1996.

Kappes, H.B., and Morewedge, C. (2016). Mental Simulation as Substitute for Experience. *Social and Personality Psychology Compass*, 10(7), 405-420.

Kauppinen, A. (2015). "What's so great about experience?" Res Philosophica 92 (2), 371–388.

Kind, A. (2020). What imagination teaches. *Becoming Someone New: Essays on Transformative Experience, Choice, and Change*, 133.

Lewis, David. 1990. "What Experience Teaches." In Mind and Cognition: A Reader, edited by William Lycan, 499–519. Oxford: Blackwell.

McCoy, J., Ullman, T., and Paul, L.A. (2019). "Modal Imagination", *Metaphysics and Cognitive Science*, eds. Alvin Goldman and Brian McLaughlin. Oxford University Press (US).

Paul, L.A. (2014). Transformative Experience. Oxford: Oxford University Press.

Paul, L. A. (2015a). Précis of Transformative Experience and reply to symposiasts Elizabeth Barnes, John Campbell, and Richard Pettigrew. Philosophy and Phenomenological Research, 760-813.

Paul, L.A. (2015b) "Transformative Choice: Discussion and replies" Res Philosophica, Vol. 92, No. 2, April 2015, pp. 473–545.

Reuter, K., & Messerli, M. (2018). Transformative decisions. *The Journal of Philosophy,* 115(6), 313-335.

Williamson, T. (2016). Knowing by Imagining. Knowledge Through Imagination, 113-123. Oxford: Oxford University Press.