



This is a repository copy of *Transformative decisions*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/127035/>

Version: Accepted Version

Article:

Reuter, K. and Messerli, M. (2018) Transformative decisions. *Journal of Philosophy*, 115 (6). pp. 313-335. ISSN 0022-362X

<https://doi.org/10.5840/jphil2018115620>

This is the peer-reviewed version of the following article: Transformative Decisions, Kevin Reuter & Michael Messerli *Journal of Philosophy* 115 (6):313-335 (2018), which has been published in final form at <https://doi.org/10.5840/jphil2018115620>. This article may be used for non-commercial purposes only.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Transformative Decisions

Kevin Reuter and Michael Messerli *

Some of the most fundamental decisions we make in our lives – such as becoming a parent or moving to a different part of the world – are transformative. According to Laurie Paul (2014), transformative decisions pose a major problem to us because they fall outside the realm of rationality. Her argument in favor of that conclusion rests on the premise that the subjective value (i.e., the value of experiencing a certain outcome of a decision) plays the central role in transformative decisions. This paper challenges that premise and hence the overall conclusion that transformative decisions usually are not rational. In the theoretical part, we specify the conditions under which transformative decisions are *possibly* rational and *likely* to be rational. The data we present in the empirical part reveal that subjective value often plays only a minor role in people’s decision-making process. Putting both parts together, we argue that people have a great chance of making rational transformative choices.

1 Introduction

Imagine that while you are on holiday, you are confronted with a choice between two Indian dishes that contain spices you have never tasted before. Suppose you are considering whether to have a child or remain childless. Alternatively, envision a slightly less mundane scenario: aliens have taken over Earth, and you have to make a choice between being transported to

*This work is fully collaborative. Please address any correspondence to kevin.reuter@philo.unibe.ch or m.messerli@sheffield.ac.uk. We would like to thank Georg Brun, Erica Cosentino, Moritz Grosse-Wentrup, Sascha Fink, Fabio Paglieri, and the audiences at workshops in Munich, London, Rome and Osnabrück for their very helpful comments on this paper. Michael Messerli was supported by a *Swiss National Science Foundation* Post-Doctoral Grant P2SKP1_171776.

an unknown planet where you can start a new life or remaining here under unknown alien rule.

In *Transformative Experience*, Paul argues that we usually cannot make a rational choice in such decision-making situations.¹ The first part of her argument rests on two premises. Premise (i) concerns the nature of the experiences that characterize those decisions – hereafter termed transformative decisions. Such decisions involve experiences that are either epistemically transformative (spices case) or both epistemically and personally transformative (child case, alien case). Having an epistemically transformative experience of x is *the only way* to know what it is like to x . For example, having the experience of being a father or eating dragon fruit is the only way to know what it is like to be a father or to eat dragon fruit. An experience that is also *personally* transformative changes one’s fundamental preferences. Experiencing life on an alien planet, for instance, would certainly change one’s core preferences.² Importantly, when a decision is epistemically transformative, the agent faces an epistemic wall that he or she cannot penetrate through imagination, testimony, or scientific investigation.³

Premise (ii) of Paul’s argument is that the *subjective value* – the value of experiencing a certain outcome of a decision – plays the central role in transformative decisions. With respect to decisions such as whether or not to have a child, Paul claims, “The value of what it is like for the agent plays the central role, if not the only role, in the decision to procreate.”⁴

¹ Laurie Paul, *Transformative Experience* (Oxford: Oxford University Press, 2014).

² Paul does not discuss purely personally transformative experiences at any length.

³ Paul reminds us that “current science doesn’t pronounce on how you’ll respond, and experiences vary widely, so you can’t rely on testimony.” (Paul, *Transformative Experience*, *op. cit.*, p. 36) Her main point is that scientific data on large groups do not give fine-grained information about how *you* respond to a particular experience (Paul, *Transformative Experience*, *op. cit.*, pp. 131-132). See Herman Cappelen and Josh Dever, “Empathy and transformative experience without the first person point of view (a reply to L.A. Paul),” *Inquiry*, LX, 3 (December 2016): 315-336 for a critical discussion about how empathy is related to the ideology of the first person.

⁴ Laurie Paul, “What You Can’t Expect When You’re Expecting,” *Res Philosophica*, XCII, 2 (April 2015): 1-23, at p. 5.

Although this aspect of the subjective value is often neglected in discussions on transformative choices, it is a crucial ingredient of Paul’s argument that the author emphasizes frequently.⁵ Consider Paul’s example in which the subjective value fails to play the central role. Many adults decide to have a child because they “need an heir, or more hands to work the farm, or whatever”.⁶ Although this is not perhaps the most common outlook among adults in our times, the general point should be clear. Paul’s challenge to the rationality of a transformative decision does not even begin to arise if people make decisions independently of the transformative experiences connected with those decisions. Thus, the challenge only applies if the subjective value plays the central role. What this means will be clarified in section 2.

Given the first two premises of the argument, we can now substantiate Paul’s claim that transformative choices cannot be rational by considering the argument in full:⁷

⁵ For example, she writes, “An essential part of your deliberation concerns what your future life will be like” (Paul, *Transformative Experience*, *op. cit.*, p. 3). For further discussion of this point, see also Paul, *Transformative Experience*, *op. cit.*, pp. 4, 25, 26, 74, 75, 85, 126.

⁶ Paul, *Transformative Experience*, *op. cit.*, p. 85.

⁷ See Jane Friedman, “Epistemically Transformative Experience,” (unpublished manuscript) for an alternative formulation of the decision problem that focuses on the role of authenticity within transformative choices. See also John Campbell, “L.A. Paul’s Transformative Experience,” *Philosophy and Phenomenological Research*, XCI, 3 (November 2015): 787-793. It seems possible, however, to formulate the challenge of rational choice without referring to the concept of authenticity.

- i. One cannot know the subjective value of at least one outcome of a transformative decision d_t .⁸
 - ii. The subjective value plays the central role in transformative decisions d_t .
 - iii. If (i) & (ii), then the ranking of options in a transformative decision d_t cannot be determined.
 - iv. If the ranking of options in a decision cannot be determined, then such a decision cannot be rational.
- \therefore Transformative decisions d_t cannot be rational.

One of Paul's most significant insights in her book is revealed by premise (iii) of the argument. If premises (i) and (ii) are both true, then it is impossible to determine an ordinal ranking of the possible options in such a decision.⁹ If the subjective value of having a child cannot be identified, and if this value plays a central role in the decision process, then there is no way to assign a specific or probable overall value to the decision to *have a child*. It follows that if no overall value can be assigned to that option, then it becomes impossible to rank the options *having a child* and *not having a child*. The final premise (iv) states a basic tenet of standard rational decision theory. A decision can be rational only if the options in a decision

⁸ A transformative decision d_t is a decision that is not merely based on a non-experiential criterion. We use this specification to exclude from the discussion those cases in which the subjective value is fully disregarded in the decision-making process.

⁹ Whereas Paul does not herself speak of ordinal ranking, we believe premise (iii) preserves the core idea of her claim. Dougherty et al. present a similar reconstruction of that premise: "If an agent cannot rationally judge the subjective value of a phenomenal outcome for her, then she cannot rationally choose between options when one of these options would lead to this phenomenal outcome." See Tom Dougherty, Sophie Horowitz and Paulina Sliwa, "Expecting the Unexpected," *Res Philosophica*, XCII, 2 (April 2015): 301-321, at p. 305.

process can be ranked.¹⁰ The conclusion is as important as it is devastating. According to Paul's argument, some of the most fundamental decisions in our life remain outside the realm of rationality. All hope is not lost, however; Paul suggests that instead of basing one's choices on the subjective value, we should make transformative decisions by asking ourselves how much we value new experiences or new selves.¹¹

In this paper, we challenge the soundness of Paul's argument by focusing primarily on premises (ii) and (iii), and we present empirical results that suggest that premise (ii) is false for many transformative decisions. At the same time, we accept Paul's arguments in favor of premises (i) and (iv). For premise (i), Paul draws an analogy with Jackson's famous Mary thought experiment: Not having seen colors so far in her life, Mary cannot know what it will feel like to perceive colors until she leaves her room. Analogously, one who has not become a parent so far in one's life cannot know what it will feel like to become a parent. As noted above, premise (iv) is part of the standard normative decision theory, according to which agents should determine the values of each possible outcome and make a choice based on the outcome with the highest expected value. Others, however, have applied different concepts of rationality. First, according to Pettigrew, Paul does not challenge the *preference-first conception*, which *merely* requires an agent to adhere to the basic axioms of decision theory, namely completeness and lack of transitivity violations.¹² Second, many theorists endorse a *satisfying theory of rationality*.

¹⁰For a discussion of ordinal preferences, see, for example, chapter 2 in Leonard Savage, *The Foundations of Statistics* (New York: John Wiley Sons Inc., 1954). Note that a decision that cannot be rational is not thereby irrational but rather falls outside the realm of rationality.

¹¹Paul, *Transformative Experience, op. cit.*, p. 123. See Eli Shupe, "Transformative Experience and the Limits of Revelation," *Philosophical Studies*, 11 (November 2016): 3119-3132 for a critical discussion of Paul's positive thesis. According to our empirical results, only in the vampire case do people primarily base their decisions on the revelatory value. Of course, the results do not tell us whether people *should* rely on the revelatory value (see sections 3 and 4).

¹²Richard Pettigrew, "Transformative Experience and Decision Theory," *Philosophy and Phenomenological Research*, XCI, 3 (November 2015): 766-774, at. p. 767. Does Petti-

Simply put, the idea behind this concept is that it is rational to choose an option in case it is sufficiently satisfactory.¹³ Of course, one might argue that in a transformative choice, an agent cannot follow the satisfying theory of rationality either because he or she cannot know whether any of the options is good enough.

Having reconstructed Paul's argument and premises, we will proceed as follows in the rest of the paper: In Section 2, we will analyze how the centrality of a decision criterion relates to the possibility of ranking options in making transformative decisions. More specifically, we will discuss what it means for the subjective value to be central such that the agent will make a choice that is *necessarily* not rational. Even if the subjective value of, for instance, becoming a parent, is not the central criterion, many transformative decisions nonetheless will fail to be rational. We therefore developed a statistical model that can tell us more precisely how likely it is that a decision is rational given the importance of the subjective value. Crucially, the likelihood increases with a decrease in the importance of the subjective value. Thus, to properly determine how likely it is that a person will make a rational transformative decision, we need to find out how people judge the importance of various decision criteria in transformative decisions, and that is exactly the task we will take up in Section 3. Through a series of studies, we found not only that premise (ii) of the argument seems mistaken but also that the data suggest

grew's challenge hold? One might at least question whether transformative decisions are *complete*. An agent might prefer neither to become a vampire nor to stay human without being indifferent to the alternatives. In this respect, there might exist a very interesting overlap with so-called hard cases of comparison. See, e.g., Ruth Chang, "Hard Choices," *APA Journal of Philosophy*, XCII, 3 (Spring 2017): 586-620); Michael Messerli and Kevin Reuter, "Hard Cases of Comparison," *Philosophical Studies*, 9 (September 2017): 2227-2250.

¹³This concept goes back to Herbert Simon, *A behavioral model of rational choice* (Santa Monica: Rand, 1953). See also Michael Weber, "A New Defense of Satisficing," in Michael Byron, ed., *Satisficing and Maximizing: Moral Theorists on Practical Reason* (Cambridge: Cambridge University Press, 2004), pp. 77-106. In Simon's view, a notion of rationality in which an agent needs to maximize utility is too demanding. That notion presupposes computational capacities that we do not possess.

that what it is like to experience a certain outcome is only one among many other decision criteria that play a role in transformative decisions. Section 4 will present a general discussion whereby we deal with possible objections to our methodology and our results.

Before we proceed, we would like to address in advance one worry that readers familiar with Paul's writings might have. The experiments we have conducted reveal which decision criteria various people in fact take into account when facing transformative decisions. However, Paul's approach is at least in part *normative* in that she is concerned with normative decision theory and not with descriptive decision theory. Are we therefore mistaken in challenging Paul's claim empirically? We are not; the normative aspects of Paul's argument concern the relation between the values of the relevant outcomes and the rationality of the decision (premise (iv)) but not which criteria are considered important for determining the value of the outcomes. Thus, we wholly agree with Paul when she states that "the normative standard for rational decision making is that the agent or decision maker *should* choose the act that has the highest expected value" (emphasis ours).¹⁴

Indeed, in Section 2 we fundamentally rely on this normative standard to develop a model for determining the relation between the centrality of a decision criterion and whether or not the decision is rational. However, Paul makes a descriptive claim about the criteria that determine which outcome has the highest value: "While the subjective values are not the only values of these outcomes, they are some of the most central and important ones, and an emphasis on them fits the dominant cultural paradigm."¹⁵ Thus, Paul's argument and the putatively dominant cultural paradigm are not about how we should determine the value of certain outcomes; rather, Paul is claiming that subjective values are most central and important. There can be little doubt, therefore, that Paul bases her normative claims on a purely descriptive

¹⁴Paul, *Transformative Experience*, *op. cit.*, p. 21.

¹⁵*Ibid.*, p. 26.

analysis of the importance of subjective values.

2 Transformative Decisions and Rational Choice

In Section 1, we started our discussion of Paul’s surprising and insightful conclusion that transformative decisions d_t are not rational by reconstructing the underlying argument in favor of that conclusion. We also highlighted that a crucial premise in her argument is the claim that the subjective values of the possible outcomes in a transformative choice are central to the decision. This claim immediately raises two questions: First, what does it mean to be central? Second, is it really the case that the subjective value plays the central role in transformative decisions? Paul does not address the first question in much detail but primarily relies on our intuitive grasp of centrality. In contrast, in regard to the second question, Paul contends that in a transformative choice such as deciding whether or not to become a parent, most people, at least in Western cultures, take the subjective value to be of utmost importance. In Section 3, we present results that strongly suggest that Paul is mistaken in her assessment of what most people consider important in transformative decisions.

In this part of the paper, we will first specify how we need to interpret centrality to validate Paul’s argument as valid (Section 2.1) and then provide a model that tells us how likely it is that a person can make a rational transformative choice given the importance she places on the subjective value (Section 2.2).

2.1 When a transformative decision is possibly rational

To better understand our investigation into the topic, consider a university hiring committee that needs to decide which candidate to hire for an open academic position. Let us look at two extremes: On the one hand, it might be the case that candidates' publication records are the only criteria taken into account. Subsequently, other factors, such as teaching skills, age, and interests, are of no concern. On the other hand, the committee might disregard the quantity and quality of the publication records and decide, for instance, to select the person who lives closest to the university. Obviously, in the former case, a candidate's publication record is all that counts and is a fortiori central to the decision-making process. In the latter case, the publication record plays no role whatsoever. The reality of most hiring committee decisions hopefully is between these two extremes: Publication records are important, but other criteria play a role, too. Analogously, on the one hand, the what-it-is-like aspect of a certain outcome might be the only criterion in a transformative decision. In that case, the decision cannot be rational (assuming the truth of the other premises in Paul's argument). On the other hand, the what-it-is-like aspect might be of no concern in transformative decisions. Accordingly, the rationality of such transformative decisions is not threatened. Again, the reality of making a transformative decision is likely to be somewhere in between these two extremes. For instance, most people making a choice about whether to become a parent will consider various decision criteria such as (a) financial situation, (b) what their partners want, and (c) whether having a child fits into their life vision and perhaps also what it will feel like to have a child. Thus, the transformative decisions people face are most likely multi-criteria decisions. Paul acknowledges situations in which the subjective value plays no role at all when, for instance, discussing the king who decides to become a parent because he needs an heir to his throne. However, Paul fails to provide an answer to the question of how important

the subjective value is when multiple criteria need to be evaluated.

An answer to this question can be given easily by making concrete what we have so far only considered in abstract terms. Take Mary, who is contemplating the choice of whether or not to become a parent. At first, Mary realizes that three criteria of varying importance are relevant to her. The most important factor to her is (a) whether her partner would like to have a baby. Less important but still of considerable weight is (b) whether she and her partner can provide financially for a child. She also (c) imagines what it will be like for her to have a baby. Although she thinks that the subjective value should be taken into account when making such a decision, the other two criteria play a greater role. We can put this in numerical terms by stating that Mary’s criteria priorities or *weights* w_i are as follows: $w_{partner} = 0.4$ (read this as saying that the criterion <partner> accounts for 40% of the decision process), $w_{costs} = 0.35$, and $w_{whatitslike} = 0.25$.¹⁶ The values for all criteria priorities should, of course, add up to 1, which is equivalent to saying that all decision criteria make up 100% of the decision process.

In the second stage of the decision process, Mary determines the so-called local alternative priorities or *values* v_i – that is, how much the three criteria favor a certain outcome.¹⁷ Let us first suppose that Mary’s partner very much would like to have a child. Thus, for criterion <partner>, Mary assigns the value 1 to the outcome “having a child” – that is, $v_{partner/pro\ child} = 1$, and 0 to “not having a child” ($v_{partner/contra\ child} = 0$). Second, Mary and her partner have a stable financial situation, although it is not ideal. Thus, for criterion

¹⁶There are at least four common assumptions made in most models of such decision processes: (1) separability of the decision criteria (to avoid double-counting), (2) numerical ascertainability of the criteria weights, (3) the weight and value attributions range between 0 and 1, and (4) for each i , $v_{outcome1} + v_{outcome2} = 1$. The empirical studies in the next section will demonstrate how we arrived at the numerical weights of each decision criterion. We will address objections to the separability of the decision criteria in the general discussion.

¹⁷This is the usual terminology used in multi-criteria decision analysis (see, e.g., Alessio Ishizaka and Philippe Nemery, *Multi-Criteria Decision Analysis: Methods and Software* (Singapore: Wiley, 2013), p. 16).

<costs>, Mary assigns the values $v_{costs/pro\ child} = 0.5$ and $v_{costs/contra\ child} = 0.5$. The subjective values $v_{whatitslike/pro\ child}$ and $v_{whatitslike/contra\ child}$ are unknown. By adding up the products of the weights and values, we can determine the overall utilities $U_{outcome} = \sum_{i=1}^n u_i = \sum_{i=1}^n v_i * w_i$ for each of the two outcomes given the first two decision criteria:

$$U_{pro\ child} = 1 * w_{partner} + 0.5 * w_{costs} = 1 * 0.4 + 0.5 * 0.35 = 0.575.$$

$$U_{contra\ child} = 0 * w_{partner} + 0.5 * w_{costs} = 0 * 0.4 + 0.5 * 0.35 = 0.175.$$

Now, even if the subjective value were to speak strongly against having a baby (e.g., $v_{whatitslike/contra\ child} = 1$ because Mary would feel completely drained by the additional responsibility and the lack of sleep), the overall utility of not having a child would be at most $U_{contra\ child} = 0.425$ ($0.175 + 0.25$) and thus still much lower than $U_{pro\ child} = 0.575$. Thus, the what-it-is-like aspect of having a child has no bearing on the outcome of Mary's decision about whether or not to have a child. In other words, the subjective value is not important enough to influence the decision process, and the transformative choice is rational. To guarantee that premise (iii) of Paul's argument is true, and hence that the transformative decision is not rational, the weight of the subjective value $w_{whatitslike}$ needs to be at least 0.5. In other words, it should account for at least 50% of the decision process. If the weight is equal to or greater than 0.5, then even if all the other criteria speak for one outcome, the overall utility can shift in favor of the opposite outcome. If the weight is lower, however, then it is possible that the transformative decision is rational. We can therefore state more precisely what is meant by *central* in premise (ii) of Paul's argument: The subjective value plays the central role in a decision process if and only if it is at least as important as all other criteria taken together.¹⁸

¹⁸We do not claim that the way we have cashed out the notion of centrality matches

How threatening are these considerations for the validity of Paul’s argument? There seem to be two possible ways to respond to the challenge we pose. First, one might maintain that the weight of $w_{whatitslike}$ is generally greater than 0.5. The empirical data in Section 3 show that such a response would be mistaken. Second, we might note that the value of 0.5 is only a cut-off point above which it is not possible to rank the outcomes ordinally. In almost all situations, however, a lower weight will still be sufficiently high, such that no rational decision will be made.

In the example above, $w_{whatitslike} = 0.25$ was not sufficiently high, but if the weight had been just a little higher, then no rational decision could have been made. To assess the strength of the second response, we now need to determine the dependency of the rationality of a transformative choice on the weights of the subjective value $w_{whatitslike}$.

2.2 When a transformative decision is likely rational

Whereas in the previous subsection we focused on the *possibility* of making a rational transformative choice, in this section we will examine how likely it is that people make rational transformative choices. In the example above, we have seen that the criteria <partner> and <costs> were too important for the subjective value to play any role.

Let us start by stating in more general terms when a transformative decision is rational. Assuming a binary choice, if one outcome is considered so much better than the other outcome, such that the subjective value does not make any difference, then the transformative decision is rational. It then simply does not matter whether the value $v_{whatitslike}$ can be determined. More precisely, we can state condition (R):

$$(R) \mid U_{outcome\ 1} - U_{outcome\ 2} \mid \geq w_{whatitslike},$$

Paul’s concept of centrality. As mentioned above, Paul remains more or less silent on this matter.

where $U_{outcome} = \sum_{i=1}^n v_i * w_i$. (R) states that the absolute difference between the overall utility of one outcome and the overall utility of the other outcome – without taking into account the subjective value – is greater than $w_{whatitslike}$. If (R) holds true, then the transformative decision is rational.

Thus, the two types of variables that are required to determine whether or not a transformative decision is rational are the weights of the criteria w_i as well as the values v_i . It seems we can now simply ask people to tell us the v_i and w_i of the various decision criteria and plug them into our formula. If Paul is correct in thinking that transformative decisions in general are not rational, we should find that most people violate condition (R) in that $w_{whatitslike}$ is sufficiently high to have a decisive influence on transformative decisions. In the next section, we present the results of several studies in which we asked hundreds of participants to report the weights w_i when confronted with various imaginary transformative decisions. However, we did not determine the values v_i of the various decision criteria because this would have required that we only consider people currently facing a transformative decision.¹⁹ The online surveys we conducted allowed us to collect the w_i relatively easily from a random sample of people from the Amazon Mechanical Turk pool. Although it would have been fascinating to interview people who are currently facing transformative decisions, this was not feasible for the purposes of this study.²⁰

Are we therefore left without a means to connect the importance of the subjective value with the rationality of a transformative decision? Fortunately, statistical models can be used to predict how likely it is that a transformative decision is rational given the weight $w_{whatitslike}$. In our first statistical model, we simplified the complexity of the situation by assigning values v_i to

¹⁹To see why, let us return to the choice of whether or not to have a child. Some decision criteria, such as <partner>, are indeterminate for most participants unless they have a partner and are currently in the process of making such a decision.

²⁰This restriction on our studies might fuel concern that the participants' ratings are not representative of decisions made in real-life contexts. We will respond to this worry in sections 3 and 4.

each criterion based on the assumption that the values v_i can be only 1 and 0. Thus, we assumed that a decision criterion either speaks fully in favor (1) of an option (e.g., having a child) or speaks wholly against (0) that option (e.g. against having a child). Besides imagining what it is like to be a parent, Mary considers two decision criteria, <partner> and <costs>, to be relevant. Thus, there are $2^2 = 4$ possibilities for how the values can be combined: (1,1), (1,0), (0,1), (0,0). Using (R) above, we can therefore calculate how likely it is that Mary will make a rational decision.

1. $|1 * 0.4 + 1 * 0.35 - (0 * 0.4 + 0 * 0.35)| = 0.75 > 0.25$ (rational)
2. $|1 * 0.4 + 0 * 0.35 - (0 * 0.4 + 1 * 0.35)| = 0.05 < 0.25$ (not rational)
3. $|0 * 0.4 + 1 * 0.35 - (1 * 0.4 + 0 * 0.35)| = 0.05 < 0.25$ (not rational)
4. $|0 * 0.4 + 0 * 0.35 - (1 * 0.4 + 1 * 0.35)| = 0.75 > 0.25$ (rational)

Thus, in two out of four (i.e., 50%) of the combinations, Mary will make a rational decision. In many transformative choices, of course, there will be many more decision criteria that might be relevant to people's choices. In all the empirical studies we conducted (see the next section), we presented people with a total of six criteria, including the subjective value. For the other five criteria, there are $2^5 = 32$ different possibilities that must be calculated to determine the overall number of rational decisions given the weights of all decision criteria, such as (1,1,1,1,1), (1,0,1,0,1).

We can now outline some specific cases related to the dependency of the rationality of a transformative decision on the subjective value of $w_{whatitslike}$. Aside from using five decision criteria (excluding <what it is like>) and normalizing the values to 1, we first consider the case in which all weights (excluding the weight of the subjective value) are equal: $w_1 = w_2 = w_3 = w_4 = w_5 = (1 - w_{whatitslike})/5$. Starting with these assumptions, we can determine the number of rational transformative choices using the following formula:

$$\text{Number of Rational Decisions} = 2 * \Theta(1 - 2 * w_{\text{whatitslike}}) + 10 * \Theta(3 - 8 * w_{\text{whatitslike}}) + 20 * \Theta(1 - 6 * w_{\text{whatitslike}})^{21}$$

The plot of the function for this very general case is depicted in Figure 1 (upper left-hand side). Importantly, the graph shows that the less important the subjective value is (the smaller $w_{\text{whatitslike}}$), the greater the number of rational decisions and hence the more likely that the person indeed will make a rational transformative decision. Given the assumptions we used, which simplified both weights w_i and values v_i , four different steps need to be distinguished. If $w_{\text{whatitslike}}$ falls below 0.5, there are two combinations for which the transformative decision is rational, namely the ones in which all other criteria speak in favor of the same outcome. However, if $w_{\text{whatitslike}}$ is less than 0.375, there are 12 combinations for which the decision turns out to be rational. A value of 0.16 or less ensures that the transformative decision is *necessarily* rational (all 32 possible combinations are rational decisions). These results might be interpreted as supporting Paul's view at least to some extent. After all, even if the subjective value does not play the central role, it might be important enough that only very few combinations turn out to be rational. We will see that this optimism is premature.

For the admittedly simple model sketched above, we find that the lower the $w_{\text{whatitslike}}$, the greater the number of rational decisions. Does this relation hold when we adopt a more realistic model? It does. To simulate a more complex situation, we dropped the assumption that all weights of nonexpe-

²¹ $\Theta(x)$ is the Heaviside step function, which takes the value 0 if $x < 0$ and 1 otherwise. We give a brief explanation of how the first term of the formula can be derived. The other terms can be calculated in a similar fashion. In the case in which all non-experiential criteria speak in favor of one outcome, then the decision is rational if $|w_1 + w_2 + w_3 + w_4 + w_5| \geq w_{\text{whatitslike}}$. We can now substitute $w_i = (1 - w_{\text{whatitslike}})/5$ and simplify the equation, which yields $1 - 2 * w_{\text{whatitslike}} > 0$. Thus, the condition is satisfied only if $w_{\text{whatitslike}}$ is less than 0.5. There are two possibilities that need to be considered: the case in which all non-experiential criteria speak in favor of one outcome and the case in which they all speak against a certain outcome. Hence, the first term is $2 * \Theta(1 - 2 * w_{\text{whatitslike}})$.

rential decision criteria are assigned the same value. Rather, we stipulated weights of increasing value: $w_1 < w_2 < w_3 < w_4 < w_5$ ²². The upper right-hand side of Figure 1 shows how the number of rational decisions depends on the importance of $w_{whatitslike}$. Again, a decrease in the importance of the subjective value leads to an increase in the likelihood of making a rational transformative choice. As the lower part of Figure 1 shows, the general increase in rational decisions with decreasing $w_{whatitslike}$ depends neither on the normalization of the values (lower left-hand side; $v_1 = 0.3, v_2 = 0.7, v_3 = 0.3, v_4 = 0.7, v_5 = 0.3$) nor on having used five additional criteria (four criteria, lower right-hand side).

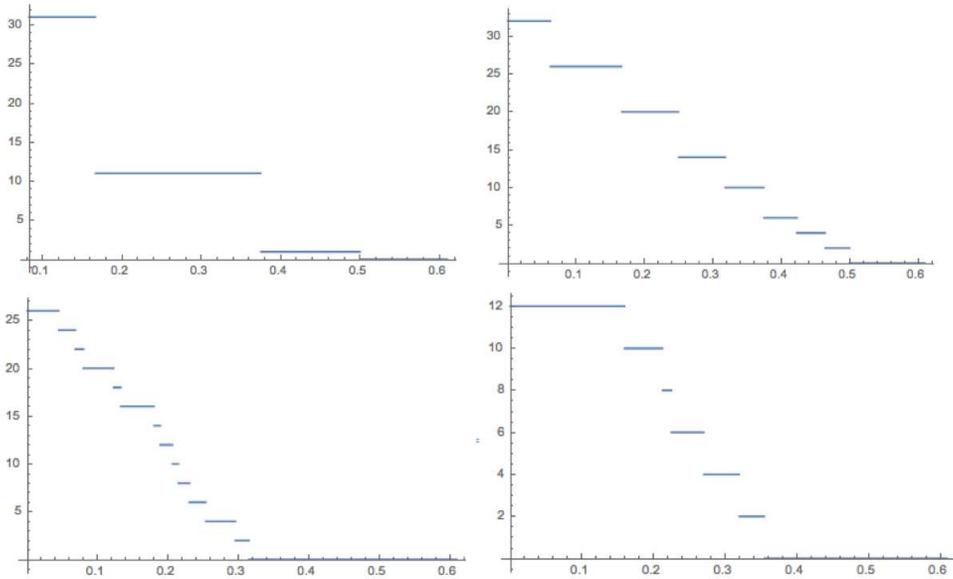


Figure 1: Plots of functions depicting the relation between the weight $w_{whatitslike}$ (x-axis) and the number of rational decisions (y-axis). In the upper left, weights w_i for all criteria except $w_{whatitslike}$ are equal. Values v_i are normalized. In the upper right, weights w_i for all criteria are different. Values v_i are normalized. In the bottom left, weights w_i for all criteria are different. Values v_i are not normalized. In the bottom right, weights w_i for all criteria are different. Only four additional criteria are used.

We varied several parameters that play a role in determining the depen-

²²More precisely, we assigned the following priorities to each criteria: $v_1 = 1 * (1 - w_{whatitslike}) / 15 < v_2 = 2 * (1 - w_{whatitslike}) / 15 < v_3 = 3 * (1 - w_{whatitslike}) / 15 < v_4 = 4 * (1 - w_{whatitslike}) / 15 < v_5 = 5 * (1 - w_{whatitslike}) / 15$.

dency of the number of rational transformative decisions on the subjective value. The results paint a clear picture: The smaller the weight of the subjective value, the greater the likelihood that a person will make a rational transformative decision. Thus, those people who do not consider the subjective value to be very important in their decision process have a high chance of making rational transformative decisions. Perhaps Paul is right, however, in arguing that a central tenet of our culture's ordinary way of thinking about transformative decisions is to place sufficiently high emphasis on the subjective value. We will now see that this view is mistaken.

3 Empirical Studies

Paul claims that transformative decisions d_t are not rational. For Paul's argument to be acceptable, the subjective value of a transformative experience needs to be the central decision criterion. In the previous section, we developed a model that allows us to calculate the likelihood that an individual will make a rational transformative choice provided we do have the weights of all the decision criteria for that individual. In this part of the paper, we present the results of an empirical investigation into transformative decisions. We examine (i) whether the subjective value of a transformative decision is central and (ii) how likely it is that subjects make rational transformative choices.

3.1 Methods

Paul and most commentators on her work tend to focus on those transformative decisions that are transformative not only epistemically (e.g., spices case) but also personally (i.e., decisions that will change the core preferences of the person making the decision). The two examples that have received the greatest attention in Paul's work are the decisions about whether to become a parent and whether to become a vampire. We also added our own case,

in which a subject must decide either to leave earth or to continue living on earth under alien rule. In what follows we present the results of three empirical studies involving those personally transformative choices.

To determine the most plausible criteria for all three cases, we ran a pretest in which we randomly assigned 81 participants to either the parent case, the vampire case, or the alien case. Subjects were asked to identify three criteria they would consider most important when making such decisions. The most common answers were then collected and categorized. In the parent case, the following five criteria were mentioned most frequently: (a) discussions with partner, (b) financial costs, (c) collecting information from books and other sources, (d) subjective value, and (e) consistency with other goals in life. Some participants also mentioned that they would contemplate how the decision would change their life. Interestingly, Paul's positive account relates closely to these responses. According to her positive view, in making transformative decisions we should base our choices on whether we might be happy to become persons with fundamentally different preferences and undergo changes to our personalities. Given participants' responses and our interest in empirically investigating Paul's suggestion, we included openness to change as a sixth decision criterion. We also would like to mention that Kauppinen has suggested that people should base their transformative choices on how consistent those choices are with previous goals.²³ Whereas Kauppinen presents a normative claim about how people should make decisions, the descriptive version of his claim closely matches criterion (e). Paul has argued repeatedly that the subjective value of an outcome of a decision needs to be understood widely as the "value of lived experience."²⁴ In the general discussion, we will explore the objection that some of the criteria that we listed as independent of the subjective value actually fall within Paul's

²³See Antti Kauppinen, "What's So Great About Experience?," *Res Philosophica*, XCII, 2 (April 2015): 371-388.

²⁴See Laurie Paul, "Transformative Choice: Discussion and Replies," *Res Philosophica*, XCII, 2 (April 2015): 473-545, at p. 513.

conception of the value of lived experience.

For the main study, a total of 181 participants were recruited on Amazon’s Mechanical Turk. Of these, 81 participants were assigned to the parent case, 39 (14 female, $M_{age} = 36.87$, $STD = 11.43$) to the vampire case, and 41 (16 female, $M_{age} = 36.61$, $STD = 12.20$) to the alien rule case. In the parent case, 28 participants had to be excluded because they were already parents. However, we did analyze their ratings to investigate the robustness of the importance of the decision criteria. Of the remaining 53 participants, 23 were female, and the mean age was $M_{age} = 31.87$ ($STD = 9.75$). The vignette in the parent case read as follows:²⁵

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

The participants were then asked to rate the importance of six randomly ordered decision criteria on a scale from 0 meaning “not at all important“ to 10 meaning “extremely important“. The decision criteria were as follows:

- I discuss it with my partner.
- I imagine what it will feel like to have the experiences and emotions when being a parent.
- I consider external factors like financial costs.
- I consider whether becoming a parent is consistent with the goals I have for my life, like my career plans and my vision of family life.
- I consider whether I will be happy to possibly become a person with fundamentally different preferences and undergo changes to my personality.

²⁵The vignette in the vampire case, for instance, was formulated following Paul’s captivating introduction (Paul, *Transformative Experience, op. cit.*, p. 1). The vignette in the alien rule scenario was as follows: Imagine that aliens have taken over the world, and you have to choose between being transported to an unknown planet where you can start a new life or staying on Earth under unknown alien rule. Now suppose you have to make the decision immediately. How would you arrive at a decision?

- I read about the pros and cons of having a baby.

The subjective value was operationalized using the criterion “I imagine what it will feel like to have the experiences and emotions when being a parent.”

3.2 Results

To evaluate the differences between the various decision criteria in relation to the rating for the subjective value, we applied a repeated measures ANOVA with *Rating* as the dependent variable and *Criteria* as the independent factor with six levels (asking partner, subjective value, cost, consistency, openness, reading). The means and *p*-values for the criteria in the parent case are listed in Table 1.

Parent				
Decision Criteria	Mean	Std Dev	Mean Difference	p-Value Comparison
Ask Partner	8.66	2.52	-0.98	0.164
Experiential Value	7.68	2.26	x	x
Costs	8.26	2.24	-0.58	1.000
Consistent	7.85	2.69	-0.17	1.000
Openness	6.43	3.04	1.25	0.174
Reading	5.45	3.34	2.23	0.001

Table 1: The importance of all six decision criteria in the parent case averaged over all participants.

For the parent case, the pairwise comparisons show that the ratings for the subjective value (7.68) were fourth after asking partner (8.66), cost (8.26), and consistency (7.85). Differences between these four criteria were not significant. The vampire scenario yielded a slightly different picture. Here, the subjective value was rated second (6.79) after openness to change (7.18). In the alien rule case, the subjective value (7.15) was the most important decision criterion but was not significantly different compared to all other decision criteria. Thus, in all three studies, the subjective value was only one of many decision criteria participants considered important.

Given the absolute ratings of each participant, we calculated the relative weights of all six criteria for each individual.²⁶ We then determined the average weights of all the criteria across all participants. When comparing the importance of these criteria, we found that the subjective value determines the outcome of the decision process only by 17%. Other criteria such as financial costs (19%) and partner’s desires (20%) seem to be on average more important to people. The results are shown in Figure 2 and demonstrate that the relative weight of the subjective value is much smaller than the cut-off value of 0.5 that we determined to be the crucial value for Paul’s argument to be sound.



Figure 2: Relative weights of all decision criteria averaged over all 52 participants in the parent case. The subjective value only received 17% of the overall weight in the decision whether or not to become a parent and is hence not more important than other criteria.

3.3 Likelihood of rational decision

Although we have established that the subjective value is not central and hence the decision process is possibly rational, we argued in section 2 that the weights themselves do not give us a final verdict on whether an agent will make a rational decision. To make that determination, we need the respective values of the decision criteria. In other words, we need to know whether a certain criterion speaks in favor or against a certain outcome. We do not know these values, primarily because they are very difficult to come by, so we developed a statistical model in section 2 that enables us to calculate the

²⁶The relative weights of the decision criteria were calculated as follows: We divided each individual rating by the sum of the ratings of all six decision criteria. Thus, imagine that a person gave the following ratings: asking partner (7), subjective value (6), costs (8), consistency (5), openness (3), and reading (1). The sum of all ratings is 30. The relative weights of the decision criteria are as follows: asking partner ($7 / 30 = 0.23$), subjective value (0.2), costs (0.27), consistency (0.17), openness (0.1), and reading (0.03). Note also that the results do not change when irrelevant criteria are added because the overall sum of the ratings is independent of the number of criteria.

likelihood that a subject will make a rational choice. By using a binary system and assigning values of either 1 or 0 to the two possible outcomes, we can specify how likely it is that the person will make a rational choice. To do this, we took the weights of each participant in our study and calculated (using a computer script) the ratio of rational and not rational choices (for details see section 2). As there are five decision criteria aside from the subjective value, the total possible combinations for each participant were 32.

Applying the statistical model revealed that only 6 out of 53 participants who completed the survey had less than a 50% chance of making a rational choice. The distribution of the number of rational decisions for each participant in the study is depicted in Figure 3. On average, the subjective value influenced people's decisions in only 35% of cases. Thus, people have approximately a 65% chance to make a rational choice. (For comparison, whereas discussions with a partner determine almost 50% of all decisions, reading literature was determinant in only 24% of the possible combinations). In the alien rule case, in which the subjective value received the highest rating, the results do not offer much stronger support for Paul's thesis because 59% of the possible decisions turned out to be rational decisions. It is therefore safe to say that given the empirical results we obtained and the statistical model we applied, it is likely that people make rational transformative decisions. We come to this conclusion not because we excluded the subjective value from the decision-making process as would the king who needs an heir; rather, we considered the subjective value, but its role can often be neglected because of its relatively low significance.

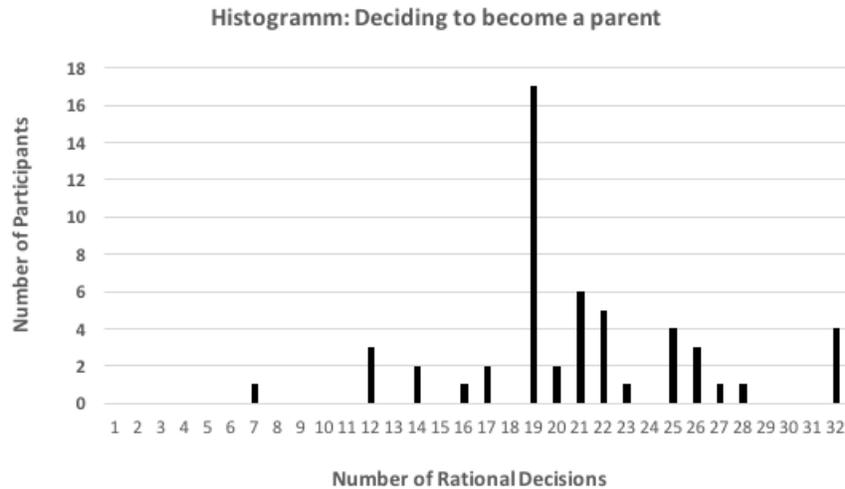


Figure 3: Histogram showing how many of a total of 32 decisions turn out to be rational for each participant using the statistical model described in the main text. For example, one participant rated the importance of the criteria such that only 7 out of 32 decisions would be rational, but for four participants, all combinations yielded rational decisions.

3.4 Discussion

The decision to become a parent is one of the most crucial decisions a person can make in his or her life. When we make this decision, we take various decision criteria under consideration, among them the subjective value of what it is like to be a parent. Other transformative decisions, such as whether or not to become a vampire, also feature largely in discussions surrounding Paul’s claim about the rationality of transformative decisions. The results of our empirical studies demonstrate that the subjective value fails to play the central role in these decision processes for most people. Instead, as the average values suggest, the subjective value is only one among many criteria that play a role in decision making. Thus, it seems that the claim that transformative decisions are not rational needs to be revised substantially. In fact, if we assume our studies are representative, most people, when facing transformative decisions, have a very good chance of making rational decisions.

4 General Discussion

Some of the most fundamental decisions we make in our lives are transformative and – if Paul’s argument is sound – fall outside the realm of rationality. Thus, the importance of a project investigating transformative decisions both empirically and analytically can hardly be underestimated. In this paper, we have accomplished three objectives: First, we highlighted a crucial premise in Paul’s argument, the investigation of which has so far been neglected. To be able to conclude that transformative decisions are not rational, the subjective value of a transformative decision needs to be *central* to the decision process. Second, we specified the conditions under which transformative decisions are *possibly* rational and *likely* to be rational. Our analysis showed that if the weight of the subjective value is lower than the weights of all other criteria taken together, the transformative decision is possibly rational. We then developed four statistical models that suggested that the lower the importance of the subjective value to the agent, the more likely it is that his or her transformative decision is rational. Third, the empirical investigation of transformative decisions revealed that the importance of the subjective value is often much lower than has so far been assumed by Laurie Paul and other scholars working on this topic. For most people, it is not only possible but also quite likely that they make rational transformative choices.

It is important to note that we have not determined how many people *actually decide* rationally when facing transformative choices. The experiments were not designed to deliver such an answer. As we argued in section 2, to make that determination one needs to collect not only the weights w_i of the various decision criteria (which we have done) but also the values v_i that indicate whether the respective criteria speak in favor of or against a certain outcome (which we have not done). Experiments that collect these values v_i are highly desirable but are also much more costly and time-consuming.²⁷

²⁷One possible exception to the claim that we do not know the direction of the criteria are the financial costs of having a child, which in our culture would mostly speak against

Given the statistical models we developed, however, it is still possible to move beyond what is actually given in the empirical data. For each and every participant in our study, the application of such models allowed us to calculate the likelihood that they would make a rational choice. The results pose a serious challenge to Paul's depiction of transformative choices. Although she is right to claim that transformative decisions are possibly not rational, the empirical reality reveals a picture in which transformative decisions are likely to be rational for the majority of people. In the remainder of this paper, we tackle five objections against our approach and results.

First, one might argue that although we tackled the question of the centrality of the subjective value, we neglected the role of possible changes in a person's core preferences, which are central to Paul's discussion. Note, however, that the issue of possible changes in a person's core preferences only becomes a problem if the decision is at least partially based on the subjective value itself. Thus, a person who makes a decision in favor of becoming a mother and brackets the transformative experience from her considerations does not make a non-rational choice even if she later regrets her decision based on her experiences.²⁸ Paul herself has stated that people who decide to become parents because, for example, they want more hands to work the farm do not fail to make rational decisions. Nonetheless, one might push this objection a little further by arguing that changes to one's preferences not only affect the values v_i we assign for each criterion to different outcomes but also the weights w_i of the various criteria themselves. Furthermore, although many participants in our studies did not rate the subjective value as sufficiently high to have an impact on their decision, the subjective value was not completely disregarded. Thus, if a person's preferences change such that the

having a child. The decision-making problem could therefore perhaps be simplified.

²⁸See Pettigrew, "Transformative Experience and Decision Theory," *op. cit.*, pp. 766-774 for an interesting discussion of the problem of weighing preferences, which might change as a result of making a choice. In this regard, an interesting question about personal identity might also exist (see Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984)), i.e., do transformative experiences undermine psychological continuity?

weight of the what-it-is-like criterion becomes so important that he or she would have decided differently, then arguably the decision was not rational. In the end, we believe this objection is also open to empirical verification. Although we do not have hard evidence for how the relative importance of decision criteria in transformative choices might change, the data we collected from people who already have children indicate that the weights of the decision criteria (at least for the parent case) are fairly robust (see the next objection).

A second objection can be mounted against the general approach that we employed in this paper. How can we be certain that the participants' responses reflect their real preferences and behaviors? Psychological studies, for instance, have shown that moral responses in controlled settings do not always correlate with real moral behavior.²⁹ However, a variety of empirical studies also have shown high consistency between people's stated preferences and their revealed preferences when it comes to aspects such as consumer choices or family planning.³⁰ Thus, the strength of this objection is likely to depend on the respective context. Unfortunately, we are not aware of any empirical evidence in favor of or against using our methodology for the investigation of transformative choices. We also are not aware of any theoretical arguments from the literature about why people's responses might not reflect their real preferences in our experimental settings. One might argue, of course, that there is a large time gap between the revealing of preferences and the actual decision making. A shorter time gap could disclose greater importance of the subjective value.

Let us give a two-pronged response. First, we can safely assume that all

²⁹Nicholas Epley and David Dunning, "Feeling "Holier Than Thou": Are Self-Serving Assessments Produced by Errors in Self- or Social Prediction?," *Journal of Personality and Social Psychology*, LXXIX, 6 (December 2000): 861-875.

³⁰Maria Loureiro, Jill McCluskey and Ron Mittelhammer, "Are Stated Preferences Good Predictors of Market Behavior?," *Land Economics*, LXXIX, 1 (February 2003): 44-55; Indralal De Silva, "Consistency between Reproductive Preferences and Behavior: The Sri Lankan Experience," *Studies in Family Planning*, XXII, 3 (May 1991): 188-197.

the participants have already made many transformative decisions – some more important, some less. It is therefore quite plausible to expect that those people can draw upon a wealth of past experiences when judging how important subjective value is for upcoming decisions. Second, for the parent case, we not only asked participants who were not yet parents but also those who were already parents. We asked the latter group to tell us how important the various criteria were when they made their decision. The results revealed remarkable consistency between non-parents and parents. The only criterion for which we recorded a significant difference was <costs>. Whereas singles tended to rate the importance of financial costs highly, parents did not believe that those costs were crucial in their decision to become parents.

The third objection targets a crucial underlying presupposition in our approach. A strong assumption of multi-criteria decision processes is that a proper evaluation of the various criteria requires that the criteria are independent of each other. This independence is important because no criterion should be double-counted. The experiments we conducted do not directly explore whether people actually make trade-offs among different criteria.³¹ To illustrate the problem, imagine that two criteria (e.g., *cost* and *consistency with previous goals*) speak against having a child. If people interpret consistency with goals as having a specific amount of money, then cost is at least partially double-counted. In other words, the criteria that speak against an outcome would then be overstated. To address this worry, we conducted a permutation test with 10.000 permutations in order to investigate whether the responses to the various criteria were indeed independent of each other.³² Two significant correlations were found: First, the criteria *ask partner* and *subjective value* were shown to be correlated ($r = 0.3623$, $p = .009$) as well as *cost* and *consistency with previous goals* ($r = 0.2821$, $p = .039$). No other significant correlations were found. While the outcome of the permutation

³¹We would like to thank an anonymous reviewer for this journal for pressing this point.

³²We are grateful to Moritz Grosse-Wentrup for helping us run the permutation test.

test demonstrates some dependency of the criterion *ask partner* on *subjective value*, the results cannot be drawn upon to question the general thrust of our empirical investigation: A correlation coefficient of 0.36 indicates a relatively weak correlation. Furthermore, no other significant correlations were found between the subjective value and any of the four other decision criteria.

Fourth, and similarly, one might challenge the idea that the importance of the what-it-is-like criteria can be determined independently from the other criteria. One might even go so far as to argue that the what-it-is-like criterion functions as a placeholder for every other criterion. For example, when you imagine how it will feel to have a child, you might consider what it will be like to spend less money during the holidays. Paul has explicitly stated that the subjective value includes “ways we’d experience ourselves in such outcomes.”³³

Kauppinen and Paul, for instance, have debated how far the criterion <consistency with goals> is an integral part of the subjective value in a transformative decision.³⁴ However, it seems that at least in the parent case the majority of the criteria, such as <asking partner>, <cost>, and <reading>, have little to no bearing on the subjective value of the outcomes. Thus, independently of what you imagine it will feel like, you either consider your partner’s opinion and the ensuing costs of the decision to be important or you do not. We also can assume that the criterion <openness to change> is distinct from the subjective value given that Paul suggests we should make transformative decisions by relying on how much we value new experiences or new selves instead of relying on the subjective value of the outcomes.

A fifth and final worry can be raised against the design of the experiments. Admittedly, we did not allow the participants to express the possible attitude that the subjective value criterion takes lexical priority over the other criteria. In other words, when we divided each individual rating (on a scale of 1–10)

³³Paul, “Transformative Choice: Discussion and Replies,” *op. cit.*, p. 514.

³⁴Kauppinen, “What’s So Great About Experience?,” *op. cit.*, pp. 371-388; Paul, “Transformative Choice: Discussion and Replies,” *op. cit.*, pp. 473-545.

by the sum of all ratings, we implicitly assumed numerical comparability. This assumption need not hold, of course. Instead, those people who rated *subjective value* the highest might have expressed that the subjective value is not only more important than any other criteria but is also lexically prior to them. To counter this objection, we would like to point out that only a single participant rated *subjective value* to be the sole, highest-rated criterion.

Empirical investigations of transformative decisions come with several methodological difficulties. We hope that further experimental studies will follow to refine the methods and hence deliver more precise results with regard to when and under which conditions people make or are likely to make rational transformative decisions. Given the results we presented in this paper, it seems highly likely that many transformative decisions fall inside the realm of rationality. We believe we have at least shifted the burden of proof to those who continue to argue for the opposite claim.

5 Conclusion

The status of transformative experiences plays a crucial role in the argument against the rationality of transformative decisions. Not only are those experiences supposed to be impenetrable, but the subjective value needs to be central in transformative decision processes. In this paper, we have analyzed the notion of centrality and have demonstrated the requirements for transformative decisions to be possibly and likely rational. We then provided an empirical investigation of transformative decisions. Perhaps surprisingly, the subjective value turned out to be not central in transformative decision processes; furthermore, the results revealed that most people are more likely than not to make rational transformative decisions even if they take into account the subjective value of the transformative choice.