

TRANSFORMATIVE CHOICE: DISCUSSION AND REPLIES

L. A. Paul

Abstract: In “What You Can’t Expect When You’re Expecting,” I argue that, if you don’t know what it’s like to be a parent, you cannot make this decision rationally—at least, not if your decision is based on what you think it would be like for you to become a parent. My argument hinges on the idea that becoming a parent is a *transformative experience*. This unique type of experience often transforms people in a deep and personal sense, and in the process, changes their preferences.

In [section 1](#), I will explain transformative experience in terms of radical first-personal epistemic and self change. In [section 2](#), I’ll explain the notion of subjective value that I use to develop the decision problem. In [section 3](#), I will discuss the way we ordinarily combine our introspective assessments with testimony and evidence. In [section 4](#), I will discuss the problems for rational decision-making. In [section 5](#), I will explore the problem of first-personally transformed future selves. In [section 6](#), I will engage with the main themes and arguments and ideas of the authors of the papers contributed to this volume.

In “What You Can’t Expect When You’re Expecting,” ([2015b](#)) I focus on a very ordinary, but deeply important, life-changing personal decision: whether to have a baby. I argue that, if you don’t know what it’s like to be a parent, you cannot make this decision rationally—at least, not if your decision is based on what you think it would be like for you to become a parent.

My argument hinges on the idea that becoming a parent is a *transformative experience*. Being a parent is a unique kind of experience that can dramatically change your core personal preferences and the nature of your lived experience. As such, it’s the kind of thing that you have to experience in order to know how it will affect you. So if you’ve never been a parent, you don’t know what it’s like to be a parent. Having the experience is necessary for you to have the capacity to represent the nature of the outcome—the lived experience of being a parent—that the subjective value of the outcome depends upon.

This unique type of experience often transforms people in a deep and personal sense, and in the process, changes their preferences. The idea isn't that you don't know what it's like to babysit, change diapers, or be very tired before you become a parent. Rather, what you don't know is its most important and distinctive feature: what it will be like to form and occupy the identity-constructing, preference-changing, physically and emotionally overwhelming perspective of being a parent.

If the salient details of the transformative experience of producing and becoming cognitively and emotionally attached to your child are epistemically inaccessible to you before you undergo this type of experience, then you cannot, from your first-personal perspective, imaginatively represent the relevant first-personal nature of the preference changes you will undergo. Because of your lack of experience, you lack the representational capacities needed to imagine, model, and grasp the nature of your future lived experience, and thus, of your future self.

In sum, you must decide whether to form yourself into a parent without knowing what it will be like to become a parent. This puts you, as a prospective parent, into a high-stakes decision problem.¹ The choice to become a parent (assuming your act is successful) is irreversible and will determine the nature of the rest of your life. Yet, the distinctive feature of being a parent, standing in a deep and loving attachment to the child you produce and raise, is epistemically inaccessible to you until you've actually become a parent. The nature and character of this attachment will, in the ordinary case, have a huge effect on your future lived experience. Thus, before you choose, you cannot assess the value of what you are likely to gain against the value of what you may lose.

This raises a special kind of philosophical problem: the choice to have a child asks you to make a decision where you must choose between earlier and later selves, with different sets of preferences, but where your earlier self lacks crucial information about the values, preferences and perspectives of your possible later selves. You cannot first-personally foresee or represent the new self you are making yourself into. We can think of the conceptual change involved as a first-personal version of a Kuhnian paradigm shift: one's first-personal view of oneself is not invariant under the epistemic transformation.²

The experience of becoming a parent is not the only kind of experience that can be transformative. It is merely an especially interesting case. It is especially interesting because many people have experienced it, the transition is reasonably abrupt, the experience can be dramatically different

¹ This is true even if you have some testimonial knowledge. While testimonial knowledge alone might license action in some low-stakes cases, it isn't enough, by itself, to license a life-changing action in a high-stakes case. (See [Moss Unpublished](#), who points out how, when an agent moves to a high-stakes context, the contents of her knowledge can license fewer actions than they did before.)

² For related comments, see [Van Fraassen 1999](#).

for different people, and an important ordinary approach to the choice (in contemporary Western society) involves deliberation and careful assessment paired with an explicit cultural narrative that urges us to “look within ourselves” to decide whether to make this major transition.

But there are other types of life transformations that can change you in this way. Going to war can be transformative (see [Zelcer 2015](#)). Developing from a child into an adult can be transformative. Descending into Alzheimer’s can be transformative. Being betrayed, or betraying someone else, can be transformative. Getting divorced can be transformative.³

A less common sort of case, but one that illustrates the idea clearly, is the dramatic life change that a congenitally blind adult would experience if he were to gain sight. In cases where we examine changes in sensory capacities, it is intuitively clear that one’s life is changed deeply and dramatically by having a distinctive new kind of sensory experience.

In all of these cases, the transformative nature of the experience can affect the real-life decisions we make when undergoing those experiences, and we must grapple with real-life philosophical issues. Once we start to look closely at major life experiences and the choices they involve, we seem to find transformation everywhere.

The wider thesis, then, is that there are distinctive philosophical issues concerning the way that we understand and construct who we are, and these issues arise in a most pressing manner when we contemplate a life-changing choice like whether to have a child. Such choices can change us deeply and permanently.

In this way, the question of whether to become a parent illustrates larger themes about the philosophical issues involved in the way we model, understand, and construct our selves. In my book, *Transformative Experience* (2014), I introduce and develop the notions of transformative experience and transformative choice and frame them in terms of transformative decision-making. There, I discuss the structure of transformative experience, and explore the tension it raises between rational decision-making and authentically forming our future selves.

In cases of transformative decision-making, you cannot grasp the subjective nature of your future lived experience, including the nature of your future self, until you become that future self, and thus you must make a life-changing decision without knowing, in the deepest sense, who you’ll become. The book develops and elaborates the structure of transformative decision-making and its implications for a wide range of big life decisions.

Many of the authors in the papers contributed to this volume raise arguments or questions about my argument, or develop the theme of transformative decision-making in interesting and novel ways. I am extremely

³ Preliminary results from psychological research done by Starmans and Bloom suggests that the transition from childhood to adulthood is transformative. Preliminary results from psychological research done by Nunziato and Cushman suggests several of the kinds of transformations I describe. I thank Christina Starmans, Paul Bloom, Josiah Nunziato and Fiery Cushman for discussion.

grateful to the contributors for their thoughtful engagement with the original argument in “What You Can’t Expect When You’re Expecting” (2015b) and with their subsequent engagement with the arguments and themes of *Transformative Experience* (2014). Below, I will attempt to address their concerns, reply to their arguments, and engage with their positive theses.

In [section 1](#), I will explain transformative experience in terms of radical first-personal epistemic and self change. In [section 2](#), I’ll explain the notion of subjective value that I use to develop the decision problem. In [section 3](#), I will explain the way life-making choices often involve assessments of one’s future lived experience in terms of personal, subjective values, and discuss the way we ordinarily combine our introspective assessments with testimony and evidence. In [section 4](#), I will explain how the ordinary, subjective deliberation involved in the decision to become a parent makes the choice transformative, and the problems this causes for rational decision-making. In [section 5](#), I will explore the problem of first-personally transformed future selves.

In [section 6](#), I will engage with the main themes and arguments and ideas of the authors of the papers contributed to this volume. [Section 6.1](#) discusses formal epistemology and decision theory and replies to John Collins, Jennifer Carr, and Thomas Dougherty, Sophie Horowitz, and Paulina Sliwa. [Section 6.2](#) discusses social choice, social justice, and social identity and replies to Rachael Briggs, Elizabeth Barnes, Rachel McKinnon, Ryan Kemp, and Muhammad Velji. [Section 6.3](#) discusses subjective value and happiness and replies to Antti Kauppinen. [Section 6.4](#) discusses decision-making in contemporary ethics, including “I’ll be glad I did it” reasoning, reasons, and self-construction, and replies to Elizabeth Harman, Dana Howard and Ruth Chang. [Section 6.5](#) discusses the problems with using contemporary empirical research to choose to have a child and replies to Nathaniel Sharadin.

1 Transformative Experience: Epistemic and Personal

An *epistemically* transformative experience is an experience that teaches you something you could not have learned without having that kind of experience. Having that experience gives you new abilities to imagine, recognize, and cognitively model possible future experiences of that kind. A *personally* transformative experience changes you in some deep and personally fundamental way, for example, by changing your core personal preferences or by changing the way you understand your desires and the kind of person you take yourself to be.

It is important to note that, as I use the phrase, a *transformative experience* is an experience that is both epistemically and personally transformative. Transformative choices and transformative decisions are choices and decisions that centrally involve transformative experiences. In many

cases, it is the degree of the epistemic transformation that creates the corresponding personal transformation—the dramatic epistemic change carries dramatic personal change along with it.

Having a child, at least in the ordinary, traditional way, involves the transformative experience of gestating, producing, and becoming attached to the child you create. In such a case, if you are a woman who has a child, you go through a distinctive and unique experience when growing, carrying, and giving birth to the child, and in the process you form a particular, distinctive and unique attachment to the actual newborn you produce. Men can go through a partly similar experience, one without the physical part of gestating and giving birth. For both parents, in the usual case, the attachment is then deepened and developed as they raise their child.

I take the experience of having a child to be unique, because physically producing a child of one's own is unlike any other kind of human experience. As a mother, in an ordinary pregnancy, you grow the child inside yourself, and produce the baby as part of the birth process. As a father, you contribute your genetic material and watch the child grow inside your partner. When a newborn is produced, both parents experience significant hormonal changes and enter new physiological states, all of which help to create the physical realizer for the intensely emotional phenomenology and cognitively rich mental states associated with the birth. These experiences contribute to the forming and strengthening of the attachment relation, and further characteristics of the nature of the attachment manifested between you and your child are determined by the particular properties of the actual child you produce. All of this generates the unique lived experience associated with having one's first child. Raising a child is then a temporally extended process that extends, deepens, and complicates this relationship.⁴

2 Subjective Values

Subjective values, as I understand them, are experientially grounded values attaching to lived experiences.⁵ These are the types of values that are involved in transformative decision-making: I describe them as “what it's like” values to emphasize that they necessarily include phenomenal value. But it is important to note that subjective values can be based on more

⁴ For simplicity, I focus on the case of parenting one's biological child. Other types of parenting can also be transformative.

⁵ To forestall confusion: to say “ x is grounded in y ” need not entail that x is entirely grounded in y . The language is similar to causal language: saying “ c causes e ” does not entail that c is the *only* cause of e . (“The striking caused the match to light” should not be understood to imply that the striking was the only cause. Oxygen in the environment, lack of extreme humidity, etc., are also causes.) So when I say that subjective values ground objective values, or that objective values depend on subjective values, this does *not* entail that subjective value is the only ground for objective value, or that objective values depend solely on subjective values.

than merely qualitative or merely sensory phenomenology: they could also include values arising from nonsensory content. Independent of any esoteric theses about qualia or phenomenology, subjective values are intended to be values that attach to the contentful features of rich, developed experiences embedded in a range of mental states such as beliefs, emotions, and desires.

Thus, as I understand it, the subjective value of a lived experience is not merely a matter of the phenomenal character of the internal characteristics of one's inner life. It's a richer value, a value that includes what it's like to live "*in this*," as John Campbell puts it (Campbell 2015; Paul 2015a). That is, it encompasses the value of what it's like to live in a particular set of circumstances, where those circumstances may include the external environment. (My reply to Kauppinen 2015 below connects with some of these issues.)

This should make it clear that subjective values are not internalist, purely qualitative values. I do not assume that the subjective value of future lived experience is determined merely by the inner, purely qualitative state of the self who is transformed by the new experience. Rather, the information gained by the discovery (even a merely qualitative discovery), functions as a necessary element in the epistemic and personal transformation of the agent. Imagining the subjective value of your future lived experience can also involve an act of *de re* imagination about the nature of your lived experience in the world.

The point of emphasizing the necessary role for experience is that you must be able to cognitively evolve yourself forward under the transformative change involved in order to prospectively grasp the subjective value of life in your possible future circumstances.⁶ What it will be like for you to live in those circumstances is informed by and infused with the qualitative information you gain when you undergo the epistemic transformation. It will have a significant effect on how your core personal preferences are formed and developed. In cases of epistemic transformation, experience is needed to teach you what your future could be like, since experience of the relevant kind is needed to give you the capacity to first-personally represent or model your possible future selves in these possible future circumstances.

There are other types of values, of course, such as impersonal moral and political values, that can also come into play when we make big decisions: I do not propose that we ignore these values. When we make big life choices, we should always make them in concert with our best objective moral, legal, and empirical standards. But for the purposes of this

⁶ These values are assigned to outcomes, because a normatively rational decision-maker is to choose to perform the act with the highest expected value, given her assignments of values to outcomes and probabilities to states. Outcomes are defined on acts by agents in states, and values for the agent attach to such outcomes. In my discussion below, I will sometimes speak loosely of assigning credences or probabilities to outcomes, where this should be understood more precisely in terms of a probability function defined on conditionals about agents performing acts in states of the world.

discussion, where we are considering an individual's personal choice, we are focusing on subjective values. That is, I am assuming there isn't some external reason that trumps or dominates your choice, making subjective deliberation irrelevant or unnecessary. As a result, the decision centrally involves your preferences concerning your future lived experiences. In the cases of interest, such preferences cannot be eliminated from your deliberations without doing violence to the natural and ordinary way you want to make the choice.

So while questions of morality and social value can apply, the context of concern for transformative decision-making is subjective lived experience.

These are the decisions I am interested in: they define a class of important and interesting life choices where a particular sort of subjective decision is called for. Not all contexts and not all transformative experiences are contexts of transformative choice, but some of the most interesting and important ones are, and they are the focus of my project. In Paul 2014, I characterized these choices as choices about who you take yourself to be and who you want to become. In what follows, I will sometimes describe these choices, understood to occur in these contexts, as *life-making choices*.

3 Subjective Deliberation

On the approach to deliberative choice I am engaging with, when making an important, life-changing decision, you want to knowledgeably assess the nature and character of the different possible outcomes of your choice so that you can choose in an informed way. More simply, you want to know what your choice means for you and your future (and for others whose futures depend on your choice). So you want to know what the possible outcomes are of your choice, and you want to know this in a way that allows you to assess and value each of them. In addition to being informed about the natures of the outcomes, you'll need to know how likely each outcome would be, given your act. Once you can assess the natures and values of the possible outcomes (and you know their likelihoods), you can determine your preferences, that is, you can determine what you prefer to have happen as the result of your choice.

Thus, to approach your life-making choice in a reflective, deliberative way, you reflect upon how you want to realize your future and then map out the options involved. You reflect on the ways you might act, on the possible outcomes of your actions, and on what those outcomes, if realized, would be like. Then you determine your preferences about how to act, given how you'd like your life to go.

What's of key interest in this discussion is that you need to be able to prospectively assess the natures of the possible outcomes of your choices in order to evaluate them properly.⁷ If you can assess the natures of these

⁷ There are important questions about likelihoods of the outcomes, or, speaking more precisely, about their credences. One important question concerns whether we have the knowledge we

possible outcomes, you can determine your preferences about your future and choose accordingly.

One especially effective way to reflect on the natures of different possible outcomes is to imaginatively project different possible futures for yourself, futures that stem from the different possible choices you could make.

Such imaginative projection is a very natural way to approach a major decision, and relies on the ability to cognitively model different possible outcomes.⁸ We do it in ordinary contexts all the time. For example, when you research possible apartments to live in, or consider buying a house, to evaluate the best place for you to live, you mentally project yourself into a future where you live in that place. The value of the outcome depends partly on how much it would cost to live there, but it also depends, importantly, on what it would be like to live there. That is, you want to assess what the lived experience would be like in that house or that apartment.

People implicitly recognize this in lots of decision contexts. In the house-hunting example, if you care about where you live and what kind of space you live in, you want to choose carefully and deliberately. Ideally you'll visit each promising place, examine it inside and out, and attempt to assess what it would be like to live there by imagining or somehow representing yourself actually living there. And if you are house hunting as part of a decision between job offers in different cities, you'll want to attempt to imagine what it would be like to live in that city, with that job, in that house. You'll try to project yourself into each possibility in order to assess and compare: should I take *this* job, and *this* apartment? Or should I choose *that* job and *that* apartment?

The more important this decision is—that is, the more important it is to you where you live and work—the more important it is that you know as much as possible about each outcome. Why? Because the better your assessment of the subjective value of the lived experience of being in that home, in that city, with that job, the better informed your decision will be.

So an important part of your deliberation involves determining the values of the possible outcomes of your action. To assess the subjective values of these possible outcomes, you need to be epistemically acquainted with them in the right way (see Lewis 1989). In particular, what is necessary for the right sort of epistemic acquaintance is that you represent the nature of the lived experience of the outcome to yourself under the subjective mode of presentation. This gives you the acquaintance you need to grasp and assess its subjective value. Arguably, the best way to manifest this representational capacity is to imaginatively represent the nature of the lived experience of the outcome. This will allow you to stand in a relation of imaginative acquaintance to that outcome, and grasp its subjective value.

need to have in order to attach credences to outcomes in a way that will license our actions. (See Moss Unpublished.) This issue will surface indirectly in a number of places, including my exchange with Dougherty et al. (2015) discussed below.

⁸ See my exchange (Paul 2015a) with Campbell (2015) for related discussion.

Is imaginative acquaintance the only way to manifest this representational capacity? I'd prefer to hold that imaginative acquaintance is the most natural and important route to assessing subjective value, but to allow for the remote possibility that there might be others. What is required for subjective valuation, however we arrive to it, is a grasp on the nature of the outcome from the subjective perspective. That is, we must represent the outcome under a suitably first-personal mode of presentation. In order to grasp the value of lived experience, we need to acquaint ourselves with it, by imaginatively representing (or perhaps by representing some other way) the nature of the experience using the subjective, or first-personal, mode of presentation.

The importance of the subjective mode of presentation is familiar from discussions in the philosophy of mind: to grasp the subjective value of an outcome involving seeing red, I need to be able to represent seeing red from my first-personal, conscious perspective. Imaginative acquaintance is the usual basis for how we represent: for example, I can represent the experience of seeing red because I am imaginatively acquainted with what it's like to see red. And, at least in ordinary contexts, experience of the right sort is necessary for us to have the ability to represent outcomes of that sort under the subjective mode of presentation.

I've been emphasizing the role of first-personal representation in assessing the subjective values of outcomes. But of course, we can get some kinds of information from other sources, such as testimony from friends and relatives and advice from experts. As you attempt to predict how you'll respond to an experience, and, correspondingly, decide how to act, you should take into account any reliable outside testimony and empirical evidence that bears on the question of what to do. In particular, you might hope to get descriptive information from these sources, such as descriptions of the various possible outcomes and other information about the numerical magnitude and valence of possible values.

We need to refine this idea in the context of this discussion, however. Recall that, in the first instance, what you want to know is the subjective values of your outcomes. Unfortunately, what testimony gives you is something different. What friends and relatives can do is describe their own experiences and outcomes. What experts can tell you is about the valence and magnitude, via numerical specifications or descriptions, of the subjective values of each outcome for the average member of a population that is relevantly similar to you. What you are getting, then, is not information that can, by itself, allow you to represent your own outcomes under the subjective mode of presentation. Rather, you are getting descriptive information about possible subjective values that is intended to aid and guide you in your representation and grasp of your own subjective values.

What you hope to do is to use this information, the general evidential facts and the testimonial evidence of those close to you, to determine what is right for you, in your situation, with regard to your life-making choice.

That is, you want to consider testimony and evidence of people similar to you when you think about how you'll respond to the experience you are considering undergoing. You'll use this to help you to know what the experience will be like for you, given what you know about the situation and what you know about yourself (which includes whatever you might learn about yourself from how others respond), so you can predict how you'll respond—that is, so you can grasp your own subjective values.

Going back to our house-hunting example: when you think about where to work and live, you should consider the testimony of friends and relatives. You should consult with others who live or who have lived in each place. Based on their experience and on what they believe about you, your sources might even give you testimony about what they think you'd prefer. You could also get official statistics about the job, about the crime rate, about the length of the commute, and about any amenities the neighborhood offers. To decide in a deliberative, informed way, you'll want to make use of these external facts and testimony, but you'll make use of them by combining them with your sense of who you are and of your own representation of how it would be for you to live and work in each place.

That is, your comparison of the subjective values of these outcomes is still based on the attempt to first-personally project yourself into each possible future. It's just that your attempt to do this should be as informed as possible by the testimony and evidence available to you.

So you weigh the testimony and evidence and consider how to apply it to your own case. Metaphorically, you survey the landscape presented by the data and testimony, and attempt to find yourself in it. You use your knowledge of how other people respond, paired with your own, first-personal assessment of who you are and how similar you take yourself to be to the others who you have data about, to prospectively assess how you think you'd respond to the experience you are considering undergoing.⁹ Formally, we might say that you use information and testimony to update your (prior) assessments of how you'll respond to the experience.

This should make it clear that taking testimony into account when you assess outcomes is not simply *replacing* what you think with what friends and family tell you. Likewise, taking scientific evidence into account doesn't mean you unreflectively *replace* how you think you'll respond in this situation with what the scientific expert tells you about how people like you tend to respond.

One of the reasons why thinking for yourself is so important in these cases is because you are making a special kind of high-stakes decision. As I have been emphasizing, transformative experiences are a distinctive kind of experience, a kind of experience that forms—or re-forms—who you are. In other words, transformative choices are life-making choices.

⁹ Of course, you may also want to predict how your act can affect others, including how it affects the lived experience of those who are close to you.

This is because transformative experiences are self-making experiences: a distinctive feature of a transformative experience is that the dramatic epistemic change involved also involves about a change in the agent's self. When making a transformative choice, it's not just about what it will be like to live somewhere new or to do something you haven't done before. It's a choice about *what it will be like to be you*. When you choose to act, in hopes of bringing about a preferred outcome, you are choosing who you'll become, based on your preferences about what self, and what sort of person, you'd like to be. As such, transformative decisions concern some of the most important and personally meaningful choices you will ever make.

This makes transformative choice into a much higher-stakes choice than the relatively mundane choice of where to live or what job to take (except to the extent that such a choice could be transformative, depending on the details). You don't just want to know what something new will be like. You want to know what *you'll* be like, that is, you want to know what sort of self you are making yourself into. And in such a high-stakes case, knowing what your future could be like before you try to make it actual is even more important. The stakes are higher, for the nature of the experience involved, the experience of being who you are, is, epistemically speaking, the most intimate sort of experience possible.

A way of emphasizing the importance of introspective reflection in these sorts of cases is to say that high-stakes choices like this can be subject to an "authenticity norm." The authenticity norm concerns the way you make life choices in concert with your first-personal understanding of who you are and what you want from life. An agent who authentically understands herself first-personally grasps her defining nature and values from the inside, that is, she knows who she is under the subjective mode of presentation. (As [Campbell \(2015\)](#) points out, we might also need to authentically understand the perspectives of others. The therapist who treats the emotional pain of her patients is able to do so authentically if she is able to imaginatively grasp salient features of their emotional experience.)

Having a first-personal grasp on the subjective values of your possible futures allows you to make choices about your future authentically. Who you take yourself to be now and whom you are making yourself into is informed by your ability to imaginatively evolve your first-personal perspective into your different possible futures. Borrowing from the philosophy of mind, we might say that your grasp on the subjective nature of your possible future lived experiences allows for an authentic mode of presentation. The idea is that, for authentic understanding, you must understand, under the subjective mode of presentation, who you are now and how you'll evolve under change into your future self. In this way, you have a grasp on what your future could be like, and on who your future self could be, so if you choose to try to bring about this particular future, you choose it authentically.

The authenticity norm comes into relief when making life-making choices, for these are high-stakes choices that we usually want to make rationally *and* authentically. Authentic rational action seems to involve acting on the basis of one's deepest principles and values, where one rationally grasps one's own principled commitments under the subjective mode of presentation.¹⁰ Such principled commitments include commitments to the kind of person you want to be and to what you care about most.

A paradigmatic case of a life-making choice is the choice to become a parent. This is a big, irreversible life choice that will probably have an outsized effect on the rest of your life. When you (carefully and deliberately) make a major life decision such as whether to become a parent, ordinarily, you think about who you are and what you want out of life, about your principles and values, and about your hopes and dreams. You look around you and see how your friends and relatives live their lives as parents—or as childfree types. You consider any available reliable testimony and empirical data about parenting and about the lives of the childfree. As you consider the relevance of the data, testimony, and any related anecdotal evidence to your value assessment, you compare yourself to the people that this evidence comes from. Finally, to determine your preferences about parenthood, you take your comparisons into account as you imaginatively consider and assess a future where you are a parent, caring for your child. To decide, you compare your expectations for your future as a parent to your expectations for a childfree life.

All of this might seem perfectly straightforward. But, as I have argued, it isn't—because the experience of becoming a parent is transformative.

4 The Transformative Choice to Become a Parent

To see how the life-making choice to become a parent is transformative, and to understand the particular challenges this raises, we need to set the decision context. First, assume that the decision maker is actually in a position to decide: that is, assume she will get pregnant and have a baby if she decides to have a child, that she has sufficient financial resources, and there are no other constraints that would prevent her from acting as she sees fit. Also assume that the child would be her first: that is, she has never experienced being a parent.

As I argued above (and in [Paul 2014; 2015b](#)), gestating, producing, and becoming attached to your child is a unique kind of experience, such that the experience of becoming a parent is epistemically transformative. The distinctive ground for the transformative nature of the experience is the epistemically distinctive attachment relation that is created between you and your child, along with its associated properties and the process that

¹⁰ So authenticity, as I understand it, involves knowledge under the subjective mode of presentation. I'd like to thank Joshua Landy for discussion.

led to its creation. This creates the special, intense, and unique feelings of parental love, care, personal engagement, and responsibility that new parents experience. These feelings deepen, develop, and change throughout the extended experience of raising your child into adulthood.

The subjective value of this distinctive type of lived experience should be understood as encompassing the external, mind-independent fact that you are standing in an attachment relation to your child. It is grounded by your loving attachment to your child, your newfound sense of responsibility and joy, as well as the rest of the features of the intense, extended lived experience of being a parent, including longer term experiences associated with raising your child to adulthood. When you discover the nature of standing in this attachment relation through the experience of being psychologically attached to the child, you gain the capacity to have the beliefs, desires and other states that define the lived experience of parenthood. In this way, you gain the capacity to represent what it's like for you to be a parent, and can thus grasp the subjective value of being a parent.

In addition to being epistemically formative, the choice to become a parent is also (usually) life changing. Who and what you care about can change, often dramatically. So if you've never parented a child, in the ordinary case, becoming a parent is both epistemically and personally transformative. The attachment changes who you take yourself to be, in the sense that you define yourself, at least partly, as the parent of your child, and changes some of your core personal preferences. Many less fundamental preference changes follow on from those.

In this way, the experience of having one's first child is what I define as a transformative experience, and the choice to have the child is a transformative choice. As I argue in [Paul 2014](#); [2015b](#), given the ordinary way we frame this choice, it is not rational. That is, on condition that we frame the choice in terms of what it would be like to become a parent, we cannot make the choice rationally.

The argument is not that there is no way to make the choice rationally. Rather, the argument is that the way many of us ordinarily *want* to make the choice, in terms of authentically grasping and assessing what it will be like to have a child *before we have a child*, is not rationally available to us.

So the problem is that, ordinarily, we want to be able to understand what it would be like to be a parent before we decide to do it, that is, we want to grasp the subjective values of our future lives as parents before we make the choice to have a child. After all, we are making a choice that will change the rest of our lives. It will have broad and significant effects on our careers, our loved ones, our financial and emotional situation, and on pretty much everything else in our lives. It's an enormous, irreversible life change, and carries with it huge responsibility and commitment. Lives depend upon it. Having a child isn't just a hobby that you spend a couple hours doing over the weekend. For many of us, it changes pretty much everything.

It's a much bigger commitment than renting an apartment for a year or taking a job in a new city. So of course, before we decide to do it, we want to know: will becoming a parent be like *this*, or will it be like *that*? Will I prefer *this* life, a life as a parent, to *that* life, a life as a childfree person? And so on. Very few people would buy a house or commit themselves to life in a new city, sight unseen, unless they had no choice. So too with parenting—we want to know what sort of life we are choosing before we undertake it. Since we cannot literally visit our possible future lives as parents before we choose (unlike the way we can visit a house, or visit a city), we try to acquaint ourselves with our possible futures in some other way. We attempt to mentally “visit” our possible futures using imaginative representation or cognitive modeling, in order to assess them and compare them, and to make an informed choice about what we want to do and who we want to become.¹¹

Once we see the importance of representing our future possibilities in order to make the choice, the nature of the problem becomes clear. Before becoming a parent, the transformative nature of the choice means that we lack the capacity to knowledgeably represent ourselves as parents under the subjective mode of presentation. As a result, we cannot grasp the subjective value of what it's like to be a parent until we actually become parents. Put in technical terms, a person who has not parented lacks a value function for the outcomes involving parenting: she cannot represent the subjective values of those outcomes.¹²

Now, a natural idea at this point is to suggest that we turn to testimony, such as the advice we get from friends and relatives and the evidence we get from scientific experts. But as I discussed above, this sort of testimony won't give you the information you need to grasp your own subjective value for being a parent, for two important reasons.

The first reason is that the testimony is merely descriptive. (Compare the attempt to know what it's like to see red merely on the basis of someone else's description of what it's like to see red. If you've never seen red, such a description won't allow you to grasp the subjective value of seeing red.) Testimony from friends, relatives, and experts might be able to give you descriptive information about the transformative experience, and about the intrinsic value of the experience. (Harman 2015 and Dougherty et al. 2015

¹¹ And of course this imaginative act is *de re*, in the sense that you are imagining what the external environment will be like, in addition to your mental life in that environment. I'd like to thank Julia Staffel for discussion (and interesting objections) about the role of imagination in decision-making.

¹² If you lack a value function then you cannot simply represent the decision problem as standard uncertainty. (See Collins 2015 and Paul 2015a for relevant discussion.) In section 4 and section 5 of this paper, I discuss (in a less formal way) the possibility of creating a replacement value function using data from experts. The first problem with this strategy is that the necessary data may not exist. The second problem is that using this replacement may violate the authenticity norm. The third problem is that the values would be for your ex post self, not your ex ante self.

argue that testimony can inform us of intrinsic value.) But it won't give you what you need to grasp the subjective value. That's partly why people tell stories about how they knew what people had told them (about parenting, or going to war, or moving to a country with a very different culture, etc.) beforehand, but there was still a distinctive and extremely important sense in which that did not prepare them for what the experience was *really* like.

Can a person at least use testimony to find out the numerical utilities of her subjective values? So, for example, even if you cannot first-personally grasp the subjective value of becoming a parent before becoming a parent, can you at least use expert testimony to know the numerical range of utilities associated with the descriptions of possible outcomes for you, along with the credences you should attach to the different propositions?

No. The trouble is that the testimony, including the evidence we've got from science, doesn't give us enough of the knowledge we need.¹³ This is our second reason for why testimony fails. The problems with incomplete evidence are not specific to the choice to have a child, though this choice provides an excellent case-study.

To start, anecdotal information from friends and relatives should always be regarded with caution. At best, it can provide some evidence of what those close to you think you should do, based on what happened to them in superficially similar situations. Scientific evidence is a better basis for rational assessments, but also has limitations, because, in the first instance, science deals in generalities concerning populations, not with specific, perfectly tailored recommendations for specific individuals in specific circumstances. (I am discussing evidence for choices made at the psychological and sociological level: other kinds of choices, such as choices made concerning biochemical outcomes at the physiological level, might be much more specific.)

That is, psychological and sociological evidence, in its current state, does not give you the individual-specific knowledge you want when making a high-stakes, life-making choice. What you want is knowledge that relates to your particular case, to your subjective values for your outcomes. But what you'll get from current psychological and sociological science is knowledge of a very general sort, concerning average effects, based on data gathered from a sample population. The problem is that the right sort of data, data that is fine tuned to your particular situation, is almost never available.¹⁴

¹³ There are subtle complications here, in part with what we take the standard for sufficient knowledge to be, and in part with respect to the interpretation of statistical information and counterfactuals. See the sections on informed consent, the fundamental identification problem, and finkish preferences in Paul 2014. Related concerns about causation, decisions, and the interpretation of evidence are discussed in Cartwright 2011.

¹⁴ That is, you want finetuned evidence and finetuned knowledge, because this is an incredibly high-stakes decision. Without evidence about your particular case, in this decision context, you lack the relevant knowledge needed for rational action. As Moss (Unpublished) puts it, in the high-stakes contexts of transformative choices, it is much harder for your probabilistic beliefs to count as knowledge (or, we might say that the weak contents of your beliefs count

The case of choosing to become a parent is an excellent example of a well-studied, widely explored life choice where, nevertheless, the science provides messy, unclear results, and gives only the grossest rough-cut estimate of the numerical utilities for any particular individual. (To get a sense of how unsatisfying it is to rely on a superficial understanding of current data to make the life-changing decision to become a parent, see [Paul 2014](#), 124–140, and my reply to [Sharadin 2015](#), below.)

The idea isn't that we can't use science to help us make decisions. Normally, we can. But that's because normally we rely on introspection about our subjective values to close the gap between the messy generalities of science and the specifics of our own personal situation.¹⁵ We finetune the general statistical information we get from science, using introspective assessments to improve our knowledge about our individual situation. But transformative experience creates a special, distinctive problem: In cases of transformative decision-making, introspective finetuning is not available, because of the epistemic inaccessibility of your future subjective values.

Compare choosing between homes in two different cities, where you have no direct acquaintance with the house or the city. You don't even have photos. You can find out what the psychological data tells you about features people tend to care about in a home or neighborhood. You can read what others say about the area. Perhaps you can also find sociological data about where people of your income, race, class, and gender prefer to live. Now assume rationally justified introspective assessment is off-limits. To choose rationally, you can only consult the psychological and sociological evidence and testimony I just described, along with any data formulated for how individuals just like you should choose, to determine your personal preferences for the outcomes.

But what data is this? Where is this trove of results that the careful, non-introspective thinker is to consult? For example, where are you supposed to find a detailed, scientifically based recommendation that, given what we know about people like you, you'd prefer to live on Main Street in the large studio with glazed windows instead of on Franklin Street in the one bedroom with a small kitchen?

Such data doesn't exist, and you can't really wait around in the hopes that it will be created. You need to decide now. So, of course, you want to go and visit each place, in order to become acquainted with each house and neighborhood. Then you can introspectively assess what you prefer and combine that with what you know from psychology and sociology. You

as knowledge but the strong (finetuned) contents don't). You need knowledge, not mere belief, to support reasons for action.

¹⁵ This may not be an especially good procedure even in contexts where transformative choice is not at issue, but it's what we do. Closing the gap successfully even in ordinary cases can be very hard, as cases involving the epistemic difficulties raised by informed consent make clear. For more on the special difficulties raised by combining transformative experience with informed consent, see the *Afterword* of [Paul 2014](#).

visit in order to improve your ability to assess the possible outcomes, and if you are taking the evidence and testimony you gathered into account (as you should), you use it to introspectively finetune the rough psychological and sociological data to fit your particular case.

If nonspecific, general scientific information without introspection isn't good enough when choosing a house, it should be obvious that it's even less acceptable for the irreversible choice to become a parent. Finetuning matters even more when making a high-stakes personal decision.

Sarah Moss ([Unpublished](#)) argues that, for probabilistic contents of beliefs, you may treat such contents as reasons for some action if and only if those contents constitute the relevant knowledge for you. As she points out, in high-stakes cases, the standard for what can license your reasons to act is correspondingly high. This is relevant to our parenting example. The incomplete, rough and general scientific evidence about how people respond to being parents doesn't give us enough knowledge to license rational action in this high-stakes choice. I need first-personal knowledge about what it would be like for me to become a parent to close the gap between the messy, general, population-level scientific evidence and knowledge of my particular subjective values. Or I need sufficiently finetuned evidence drawn from a more complete science, evidence that can support probabilistic knowledge about my individual case.

The problems for transformative decision-making based on scientific evidence, thus far, have been practical.¹⁶ But there is another sense in which science cannot give you the knowledge you need to rationally motivate a life-changing transformative decision. In cases of transformative decision-making, we have an in-principle problem with *ex ante* decisions made for *ex post* subjective values and preferences.¹⁷ This problem is both a problem for decision theory in its own right, as it involves the possibility of incommensurable preferences across selves, as well as a problem involving authenticity, because authentic choice-making requires the right sort of knowledge about who you are making yourself into.

Another way we might think of the problem is in terms of motivation: without first-personal insight into the self I'll become, how can I be motivated to become that future self, if that self is incommensurable with who I

¹⁶ Strictly speaking, the problem of the generalities of science isn't merely practical, because empirical evidence concerns average effects for members of populations, not individual effects. This (and related concerns, such as the fundamental identification problem) might make it in-principle impossible to get sufficiently finetuned evidence in these high-stakes cases, making it in-principle impossible to get the kind of knowledge we'd need to license action in transformative decision contexts. Here, it depends on just how high the stakes are and on just how specific one can be about the relevant population. See [Paul 2014](#) and [Cartwright 2011](#) for related discussion.

¹⁷ See [Pettigrew 2015](#) and [Paul 2015a](#) for further discussion.

am now?¹⁸ I'll discuss this problem briefly in [section 5](#), below, and it will surface in many of my replies in [section 6](#).

5 Rational Decision-making Under Radical Change

The case of choosing to become a parent illustrates how, if you approach a life-making transformative decision intending to assess, understand, and then choose between the different ways your future could be like, you cannot make the choice rationally. This is because of the combination of epistemic and personal change involved in the transformative experience. The epistemic problem arises because the decision is to be made based on a subjectively informed assessment of your possible futures. The personal problem arises because the decision involves the possibility of undergoing an experience that changes you from the self you are now into a different, new self. Together, these problems create a situation where the nature of the experience and the way you'll respond to it involves the possibility of change from one self into an incommensurable, epistemically inaccessible self.¹⁹

The epistemic situation is analogous to problems with incommensurable theoretical paradigms and the rationality of discovery and theory change from one scientific paradigm to the next. In a life-changing choice, you, the chooser, cannot escape your current perspective. In a Kuhnian sense, you are trapped within the "normal" paradigm of your current self. When confronted with a transformative choice, you must decide whether to replace your current self and its perspective with a new self and that self's perspective. Yet, to grasp the nature of the new self you could become, you must undergo the transformative change, because the nature of that future self is epistemically inaccessible to you before the transformation.

There is no problem of strict personal identity here: we can assume both selves are metaphysically the same person. The trouble is that they are psychologically incommensurable with each other. And so a kind of existential crisis arises: due to the epistemic inaccessibility of the future self's perspective from the current self's perspective, the agent cannot know who she is making herself into. Moreover, due to the incommensurability of the preference change, she cannot adopt a principled decision rule to prefer one set of preferences over the other. Because she cannot step into a neutral first-personal perspective in order to evaluate and compare each possible

¹⁸ Moss (Unpublished, 9.6) explores intuitive ways to understand what contents may count as your *motivating* or *personal* reasons for action. I think this relates to what I described above as the authenticity norm, where there is an intuitive sense in which your personal reasons for action must be contents that you can represent. (This might connect to questions explored by time-slice epistemologists. If at *t1* the agent cannot first-personally represent the self she'd be at *t2*, why should she be motivated to become that self?)

¹⁹ See my exchange with Collins 2015 for related discussion.

successor first-personal perspective, she cannot formulate a higher-order rule that will adjudicate the decision for her.²⁰

One way to represent this problem from a decision-theoretic perspective is to argue that, before your choice, if you cannot represent (or “see”) the outcomes in the necessary way, you lack psychologically real, rationally assignable preferences about your post-transformation outcomes. We can describe this as a violation of a standard axiom of decision theory, the completeness axiom.²¹ Completeness requires agents to have definite preferences for any gamble, such that the agent rationally prefers to take the gamble, or disprefers it, or is indifferent towards it. (We might think, loosely, of the different selves an agent could become as the different possible “prizes” she might win from a gamble.) On an account where we understand the agent’s preferences to be psychologically real, an agent who cannot have rational preferences about her post-transformation selves lacks the needed attitudes, and so the axiom is violated. This is simply another way to put my point that an agent without a value function for transformative outcomes is an agent without a (standard) model for a rational decision.²²

The possibility that the transformative change involves incommensurable self-perspectives and the replacement of one’s current self with an epistemically inaccessible, incommensurable future self is what creates an in-principle problem for using testimony from others and for using future values determined by experts. (Note: this is different from the practical problem I discussed in [section 4](#), the problem that, given the state of current psychology and sociology, relying on testimony and evidence won’t get you sufficiently finetuned knowledge.)

The rabid fan of testimony will advise you to dispense with introspection. You are to make your decision entirely on the basis of what others tell you about their experiences, or about what scientists have discovered about others who have gone through similar experiences.

But ignoring introspection has serious costs. The first cost involves deciding without knowledge of the value of the lived experience. As John Collins argues in his contribution to this volume, an agent might be rationally neophobic, that is, she might be rationally averse to making a decision without grasping the subjective values of the outcomes she could bring about. In high-stakes cases, it’s especially plausible that we’d be rationally averse to making life-changing decisions without knowing the subjective values of our possible futures. Such aversion also has connections to the desirability of authenticity and the role of motivation in personal decision-making.

²⁰ The source of this problem is the fact that the agent only learns what she needs to know, that is, she only has access to her post-transformative perspective, *after* the transformative experience. See [Pettigrew 2015](#) and [Paul 2015a](#) for discussion.

²¹ I thank Alan Hájek for suggesting this way of framing the discussion.

²² For related discussion, see [Collins 2015](#) and my reply below.

The second cost stems from the fact that you are to use data and testimony to determine your *ex post* values. The trouble with such testimony is that it is given by those who have undergone the experience, and so concerns the *ex post* self, the self that would result from undergoing the experience. That is, it's relevant to who you'll become as the result of the transformative experience, and purports to tell you how you can expect you'd assign values to the outcomes at that time. But from your *ex ante* perspective, that is, from the perspective of the self who makes the decision, given the transformative nature of the experience, your future self could be incommensurable with your current self. In other words, the subjective values given by testimony for the outcomes are values of your merely possible future selves, not your current self, the self who is making the decision. And the trouble is that, when you face a transformative choice, even if expert testimony can tell you what to choose, you still face an existential problem: Will you really be happier after the transformative change—or will you just be a different self?²³

This matters, because you might find some of your possible future selves epistemically alien to who you are now. For example, you might be a career-driven childfree person who finds small children annoying, but if you went through the transformative experience of becoming a parent, you'd enjoy spending time around babies and look forward to your hours at the playground. As your current, childfree self, you might find such a future self repulsive—she is deeply epistemically alien to who you are now. If so, why should the subjective values of that merely possible future self be relevant to who you are now?²⁴ Put another way, even if we assume that scientific data and testimony could provide you with your future self's values, that is, with your *ex post* values after the transformation, this isn't enough. If, at *t1*, you want to choose consistent with your current self's preferences, and if she prefers to remain who she is, what matters for the rationality of your choice at *t1* are what you determine *ex ante* about the values for the self at *t2*, not what the *ex post* values are for the self at *t2*.

Normally, as I discussed in [section 3](#), to assess our values for a future lived experience, we imaginatively project ourselves into the future possibility, and prospectively assess it to determine, *ex ante*, our *ex post* values. This works if there is no transformation of the self, and it might even work if you could assess the transformation from a “self-neutral” perspective.²⁵ But the combination of epistemic and personal transformation in transformative choice makes this impossible. You cannot imaginatively project yourself “through” the experience to assess your future self's first-personal

²³ Again, we can see this as a problem with completeness. It can also be seen as a problem with van Fraassen's Reflection Principle. In [Paul 2014](#), I discuss the problem in the Afterword, in the section “Finkish Preferences.”

²⁴ For related discussion, see [Barnes 2015b](#), [Briggs 2015](#) (and my reply below), and [Paul 2015a](#).

²⁵ Compare Chang's (2015) suggestion that we adopt a “master utility” function in this situation.

perspective, for it is inaccessible to you, and thus you cannot prospectively model your transformed first-personal perspective in that possible outcome. As a result, you risk forming yourself into an alien self when you undergo a transformative experience. Relying solely on data or testimony to tell you your future values obscures this fact, because it merely tells you the values of your transformed self, not the values for who you are now. I discuss this cost further in my reply to Briggs, my reply to Dougherty, Horowitz and Sliwa, my reply to Harman, and my reply to Chang.

The problem of inaccessible, alien future selves also brings out how exploring transformative experience in the context of a psychologically rich decision theory or a formal epistemology gives us a richer perspective on the way that rationality and authenticity are intertwined with transformative choice. To act authentically, you must be adequately informed of the nature of the outcomes of your choice: authenticity can require knowledge under the subjective mode of presentation. Moreover, consistent with more traditional notions of authenticity, your values should be formed by who you are and what you know about yourself, and by your core principles and commitments. Simply adopting values given to you by external authorities is inauthentic. The same holds for moral values: you need to understand and grasp these values for yourself in order for them to be *your* moral values.

There *are* ways to rationally and authentically prefer to disprefer your current values, for example, when your higher-order preferences are to change your current higher-order preference structure. For such a decision to be authentic, you must understand that you are choosing to discover a new self, that is, you must reflectively decide to replace your current self with the new self that you will discover, knowing that you don't know what your future self will be like. (In an important sense, you prefer to annihilate your current self.) In *Transformative Experience* (2014), I argue that this preference structure involves a preference for discovery and revelation, and may be able to resolve the tension between rationality and authenticity in some transformative choices.

Once we see how epistemic and personal transformation work, it becomes apparent that some of life's biggest decisions are life-making choices. The deep philosophical problem is that these life-making choices involve experiences that teach us things we cannot first-personally know about from any other source but the experience itself. With many life-making choices, we only learn what we need to know after we've done it, and we change in the process of doing it.

The lesson I draw is that an approach to life that is both rational and authentic requires epistemic humility: life may be more about discovery and coming to terms with who we've made ourselves into than about carefully executing a plan of self-realization.

6 Discussion and Replies

6.1 Formal Epistemology and Decision Theory

John Collins (2015) argues that decision theory needs to make conceptual room for the rationality of neophobia, that is, for a rational aversion to the new and unfamiliar—just as we need to make room for psychological facts about agents' attitudes concerning risk or ambiguity.

He explores how, for a decision concerning an epistemically (but not personally) transformative experience, we can model the choice using a familiar proxy as a synthetic lottery, and represent an agent's preference for the synthetic lottery as a neophobic preference. We might then develop a decision rule for such preferences. I find his positive proposal and diagnosis of the source of neophobic preference structures creative and insightful, although I don't think that the deep source of neophobia is indeterminacy. But despite our differences, I completely agree that we need to make room in decision theory for neophobic and neophilic preference structures, and find his decision rule for neophobic agents compelling.

Collins introduces his argument with a discussion of the metaphysics of utilities, arguing that I am committed to a species of nonconstructive realism for utilities. Constructive realists take an agent's utilities to be metaphysically constituted by her preferences. Nonconstructive realists deny this. Since I am not committed to nonconstructive realism, I'll start with a brief discussion of his claim and then move to a positive discussion of his substantive proposal.

I take preferences to be psychologically real, but I have no stronger claim about the metaphysics of utilities. On my view, an agent facing an epistemically transformative experience lacks psychologically real, rationally assignable utilities concerning the epistemically inaccessible outcomes because she cannot represent (or "see") the outcomes in the necessary way.²⁶ If utilities are understood in constructive realist terms, then on this view, psychologically real preferences don't exist either. The agent lacks the capacity to grasp or entertain the natures of the relevant outcomes, and thus lacks the desires and beliefs needed for her to have psychologically real preferences.

If an agent lacks the ability to have psychologically real preferences for a decision situation, we find ourselves without a model for rational decision in that situation. That is, we lack the capacity to represent and model the choice "in the usual and obvious way, as a gamble that might yield any one of a range of possible utility values, depending on how things turn out to be" (2015, 287).²⁷

²⁶ Here I adopt a nice locution from Carr 2015.

²⁷ None of this rules out the possibility that the agent could have preferences based on confusion, or on false, incorrect, or otherwise inappropriate beliefs. Such preferences won't help us with rational decision-making in this context.

So the nature of the problem with epistemically transformative experience does not depend on whether one is a constructive or a nonconstructive realist, for an agent without rational preferences is also an agent without a rational decision model. But no matter: I agree with Collins's main thesis, that decision theory needs to make conceptual room for neophobia.²⁸

Collins proposes that we explore the contours of the problem of the inaccessibility of the epistemically unknown outcomes using a synthetic lottery as a replacement model.

- (1) For any possible utility value x that the epistemically transformative experience may turn out to have for the agent, there is a possible outcome to the lottery that is both (a) experientially familiar to the agent and (b) has a utility that is (arbitrarily) close to x .
- (2) The chances of the various possible outcomes to the lottery are weighted so as to correspond to the agent's subjective probability distribution over the range of possible utilities that the epistemically transformative option A may turn out to have, whatever that subjective probability distribution happens to be. (290)

Perhaps an agent can determine her expected utilities using a synthetic lottery over the familiar space of options, then map that choice onto the space of epistemically transformative outcomes to determine her expected utilities for the transformative decision. Then, even if we can't grasp the characters of the unknown outcomes in order to determine their utilities and the expected utility of possible acts involving them, we can construct a proxy model with the same structure, based on mapping known utilities for known options to unknown utilities for ungraspable options in order to use it as a proxy for the model we'd have constructed if we were able to explore the space of options in the usual way.

This is an extremely interesting suggestion, and I discuss its implications in detail in [Paul Unpublished](#). Of course, we'd need to know how to construct the mapping from the space represented by the proxy model to the space of the epistemically transformative situation, but let's assume that this is discoverable empirically or knowable via testimony from others.

As I argued above, the need for the model doesn't turn on whether one is a constructive or nonconstructive realist, for if one is a constructive realist the proposal is just a little more radical: instead of taking the values of the outcomes as synthetic replacements, take the entire model, the synthetic lottery itself, as a synthetic replacement for the epistemically inaccessible model of the decision. In constructive realist terms, we can think of the

²⁸ This small dispute is probably due to the fact that Collins was working from a draft copy of my book manuscript rather than the final version. In the draft copy I did not specify my preferred metaphysics of preferences and utilities.

synthetic lottery as providing us with synthetic preferences from which we construct synthetic utilities.

The deeper and more interesting issue here is whether there is a dimension of value that is not captured by the numerical values for utilities (whether or not they are constructed from preferences) that are assigned to decision outcomes via testimony. We can understand neophobia as arising from a situation where we cannot first-personally grasp or entertain the qualitative nature of the lived experiences described by the outcomes in the model. Even if someone else can tell me what the numerical values of my utilities are for outcomes involving durian-eating, there's something important about these outcomes that mere description or testimony can't communicate. Knowledge of numerical values for utilities alone isn't sufficient.

We might diagnose the source of the problem as one with our inability to form preferences, or to grasp or entertain the subjective values of these outcomes, as I do. (For more on these subjective values, see my discussion in [section 2](#).) Or we might diagnose it as an inability to resolve an indeterminacy about the possible utilities I might experience.²⁹

Collins argues that the best response for the defender of orthodox decision theory is to regard the decision problem with epistemically transformative outcomes as a problem involving a "basic and irresolvable" indeterminacy about utilities.³⁰ He writes:

That's why the orthodox decision theorist's suggestion that we elicit her utility for the transformative outcome by the method of constructing a synthetic lottery need not always work. It's not possible to elicit a sharp determinate value for the utility of an outcome when it is just a fact that no such unique value exists. The synthetic lottery may yield some unique number, but so what? It's providing an answer to a different question. (2015, 287)

He suggests that the irresolvable indeterminacy could stem from the agent's inability to assess and compare the values of different outcomes, given her lack of acquaintance with them.

He then argues that, nevertheless, there is a rational basis for the neophobic agent to opt for a synthetic lottery over a lottery with unknown outcomes, drawing on Isaac Levi's work on the Allais problem. I find his proposal interesting, and it should certainly be included in the set of options we should consult when faced with a transformative decision problem.

²⁹ Collins, in conversation, points out that he and I might not be at odds here. If we both endorse constructive realism about utility then these would be two ways of saying the same thing: indeterminacy about utility would arise out of the inability to form psychologically real preferences.

³⁰ See [Barnes 2014](#) for an excellent discussion and defense of fundamental metaphysical indeterminacy.

However, it is important to see that, whether we take the source of neophobia to be our inability to grasp subjective values, or we take it to be our inability to resolve indeterminacies about utilities and employ synthetic lotteries, in transformative decision contexts, the fundamental tension between rationality and authenticity remains.

In fact, Collins's proposal for modeling epistemically transformative decisions gives us a new and especially clear way to bring out this fundamental tension. We can see this by exploring how a synthetic model might be used to make a decision that is both epistemically *and personally* transformative.

Recall that the synthetic lottery works by substituting unfamiliar outcomes with familiar outcomes with the same utility. Now consider using a synthetic lottery for a transformative choice, like the choice to become a parent. If I'm neophobic, then even a perfectly constructed synthetic lottery will fail to yield an expected utility relevant to my decision problem.

In this case, I'm to use a synthetic lottery to decide whether to become a parent.

For example, let's say I've already had the transformative experience of becoming a doctor, and so I am familiar with the kind of epistemic and personal transformation involved in such a life change. Now I am considering becoming a parent, and I'm told that the possible outcomes of this new life change can be assigned utilities that mirror, in the right way, the values of the possible outcomes associated with my becoming a doctor. So I can use the synthetic lottery to determine that, if I maximized my utility by becoming a doctor, I will maximize my utility by becoming a parent.

The important detail in this case is that I'm using the synthetic lottery to decide whether to undergo an experience that could result in an alien future self, a self whose first-personal perspective is epistemically inaccessible to me now. What's familiar to me is the dramatic personal change involved with the experience of becoming a doctor. But for just that reason, I know that the utilities attached to the outcomes of becoming a parent carry with them the possibility that they are utilities for a self who is psychologically alien to who I am now. My choice is whether to make myself into this future self.

We might, then, understand neophobia as a perfectly rational aversion to the possibility of becoming an unknown self—and neophilia might be understood as a kind of self-hatred, that is, as a desire to replace one's known self with an unknown self. And so the central tension between rationality and authenticity remains.³¹

³¹ I might use the synthetic lottery as a model for a decision between revelation and the status quo, where I am choosing to discover what it's like to become a new kind of person. See Paul 2014, Ch. 4, "The shock of the new." What the synthetic lottery can capture here is what the agent is familiar with, that is, what it's like to become a new kind of person, and whether this is desirable.

Like Collins, Jennifer Carr (2015) focuses on epistemically transformative experiences. She develops a model for epistemic transformation in epistemic decision theory. As Carr puts it, the question is:

How can there be a decision theory for partial credence functions, when decisions hinge on possibilities the agent can't entertain? The problem is not uncertainty: it's not simply that the agent is unsure of the outcomes of her actions. Rather, the problem is limited conceptual resources: there are some possibilities that the agent can't "see," propositions she isn't in a position to entertain. (217)

An epistemic decision theory can be used to develop and model normative epistemic facts about agents, and a rational epistemic decision model can be understood as a model for epistemically rational belief and behavior.

Epistemic utilities (values) are understood in terms of what is epistemically desirable or in terms of epistemic goods, and credences are strengths of belief. The idea is to give epistemology an overtly decision-theoretic framing in order to discover and represent epistemological norms for rational agents, such as which epistemic states they should adopt given their evidence.

In Paul 2014 and 2015b, I formulate the questions about transformative decision-making neutrally, but they can be explored within more specific frameworks, like that of epistemic decision theory. The problem that transformative experience raises for epistemic decision theory concerns our epistemic guide for life-determining and life-changing choices, that is, our epistemic norms for how we should believe and act in order to maximize our epistemic utility when we make big life decisions.

Epistemically transformative experience, viewed through the lens of epistemic decision theory, generates a problem. How? Because the transformative nature of the experience rules out the possibility that the agent can subjectively grasp her utilities, since she cannot "see" these utilities or entertain the relevant propositions.

A solution, of course, is for an expert to tell the agent what her utilities are. However, as I discussed above, in most cases, such expert advice doesn't exist. And even if it does, simply knowing the numerical assignments for outcomes via testimony is not sufficient for an agent to fully grasp the nature of the outcomes, and thus is not sufficient for her to determine her expected subjective value. There's a dimension of the subjective value that can only be grasped by having the right sort of experience. It's an open question whether this arises from the subjective value s for outcome o not being representable by numerical value n , or whether it arises from the psychological fact that the agent cannot grasp or understand everything represented by n without the requisite experience.

The central tension between rationality and authenticity, then, can be illustrated in epistemic decision theoretic terms. The relevant norm of

epistemic rationality requires an agent to act in accordance with maximizing her expected epistemic utility. The relevant (cultural or practical) norm for authentic choice and action defines an agent as informed only if she can first-personally grasp, understand and subjectively value her outcomes. This norm of authenticity holds that agents should make life-defining choices as informed agents, which means an agent should knowingly make her life choices partly by understanding who she is, who she wants to become, and what her lived experience and the lived experiences of others could be like as the result of her choices.

In transformative decisions, if the agent's expected utility is only available to her via expert testimony, then to meet the epistemic norm the agent must make a decision about her future without a first-personal grasp on her future subjective values. To act (epistemically) rationally, the agent must violate the authenticity norm.³²

As the clash between norms suggests, the most interesting version of the problem involves experience that is both epistemically and personally transformative. Such experience concerns big life decisions, and highlights what is really at stake in this discussion, for it highlights how an agent's ability to grasp the first-personal perspective of her future self is a central part of her ability to grasp who she is making herself into. The problem is that knowing the numerical utility for the self who results from a transformative personal experience is not sufficient for an agent to know who that self is, and whether that future self is desirable from her current point of view, for the personally transformative experience is *also* epistemically transformative. As a result, the fundamental tension between epistemic rationality and authenticity is highest in the cases that matter the most from a real world perspective.

Carr's approach to the issues involving epistemically transformative experience and epistemic decision theory focuses on how we are to understand the way agents with partial credence functions are to expand their views when they experience an epistemic transformation. How are we to understand the epistemic norms for an agent who discovers new propositions or can see new possibilities? She frames the epistemic transformation in terms of the agent's credence function changing domains from one set of propositions to a new, expanded set, and defends the view that credence functions with different domains are sometimes comparable.

On Carr's view, we can think of epistemic decision theory as including normative epistemic constraints, defined by the value of conceptual resources, on credence functions. In her paper, Carr focuses on accuracy-first epistemology, and explores different constraints under which credence functions with different domains are comparable.

³² Interestingly, perhaps the best way for the agent to respond to her situation is to authentically grasp that she doesn't know her future subjective values and choose with respect to her desire for revelation. See my chapter 4, "The shock of the new," for discussion of this idea, as well [Campbell 2015](#) and [Paul 2015a](#).

I found her preferred proposal for modeling epistemic transformation, where we regard nonattitudes towards propositions as having the same utility as having maximally unopinionated attitudes to those propositions, to be an interesting and important way to understand some of the epistemological implications of epistemic transformation. The question I want to pose for her, however, asks how such an approach affects the way we should understand transformative decision-making that involves personal change.

In particular, how does the proposal pan out in a context of epistemic personal discovery, that is, a context where the agent discovers a new way to conceive of herself, and of the world in relation to herself? As I noted above, one of the central puzzles of transformative decision-making concerns decisions involving epistemically and personally transformative experiences. The puzzle arises when an agent's ability to see new propositions generates new, incommensurable personal preferences. How should we regard the utility of remaining maximally unopinionated in this context?

Moreover, what are the constraints on epistemic expansion when an agent extends her conceptual grasp on the world in a way that gives her a new, incommensurable representation of herself, or of the way the world is? What are the constraints on epistemic expansion when an agent changes (or an agent contracts) her conceptual grasp on herself and the world in a way that gives her a new, incommensurable representations of them? Such epistemic change, tied with personal change, might be more complex than mere contraction or mere expansion, if, as seems likely, some propositions are simply replaced by others.

Carr's proposal explores foundational questions about the epistemology of epistemic discovery and change. It would be good to know more about how it bears on more complicated cases involving change that is at once epistemic and personal.

In their engaging paper, Tom Dougherty, Sophie Horowitz, and Paulina Sliwa (2015) (hereafter, "DHS") defend the decision-theoretic homeland by raising a series of objections to my argument that the phenomenon of epistemically transformative experiences contributes to a new and distinctive problem for decision-making. Their central claim is that epistemically transformative experience "does no special work" because it is possible to estimate the intrinsic value of outcomes that involve epistemically transformative experiences even if one has not had that type of experience. They then explore ways of framing transformative decision-making as involving the kind of uncertainty that can be represented by imprecise credences. I found their discussion useful and informative in its exploration of some of the controversial features of the debate over transformative experience. However, while their arguments raise thoughtful objections that deserve attention, in the end, I reject their central claim, arguing that they have conflated subjective value with a different notion of value, intrinsic value, and I deny that the epistemic problems raised by transformative decision-making

are simply epistemic problems that can arise in cases where subjective experiences do not feature.

DHS target the following premise:

(Premise 3) If an agent does not know what it is like to have an experience, and this experience is constitutive of a “phenomenal outcome,” then she cannot rationally judge the subjective value of this outcome for her.

Premise 3 isn’t a premise I actually defend: it is a premise culled from their characterization of my view. However, I shall accept it for the purposes of our discussion.

DHS reject Premise 3. They argue that distinguishing phenomenal character from value

allows us to also draw an *epistemological distinction* between awareness of an experience’s phenomenal character and awareness of its value. . . . Once drawn, this epistemological distinction should make us suspicious of Premise 3. From the fact that an experience is epistemically transformative, it only follows that the agent is not antecedently in a position to know what the experience would be like. This is consistent with the agent being able to rationally estimate the experience’s value. (307)

However, distinguishing between phenomenal character and value won’t do the work they want it to do. I agree that we can make an epistemological distinction between an agent’s awareness of an experience’s phenomenal character and her awareness of its subjective value. But this does not undermine the fact that, in the cases we are examining, experience is necessary for an agent to grasp and represent the subjective value of the outcome.

Recall that the subjective value of an outcome (a future experiential state of an agent) ontologically depends on the nature of the lived experience that constitutes it. We can certainly distinguish between an agent’s being aware of, or grasping, the nature and character of the lived experience that constitutes the outcome, and the agent’s being aware of, or grasping, the subjective value that depends on the nature and character of the lived experience. But the agent’s inability to grasp and represent the propositions concerning the nature and character of the outcome, or, we might say, her inability to grasp these propositions under the right mode of presentation, implies that she cannot be aware of the subjective value of the outcome.

And, of course, subjective value is precisely the sort of value we are concerned with in this discussion. As I put it in [Paul 2015a](#),³³ for many

³³ Another interesting distinction we can make involves the distinction between knowing the numerical value that an expert tells her to assign to the subjective value and the epistemic act of grasping or being aware of the subjective value.

life-changing decisions, the agent wants to assess her options by assessing the subjective value of her possible future lived experiences. Ideally, her assessment involves a determination of the subjective value of each possible outcome of her decision, that is, the subjective value of each possible lived experience, by imaginatively grasping what it would be like for her to live in those possible circumstances. This does not imply that subjective value merely concerns the character of one's internal mental life, or that it is somehow merely self-interested. Subjective value, instead, is concerned with the nature and character of an agent's lived experience, which can include her experience of her environment (and can include assessments of subjective values for other people as well). (For an expanded defense of the importance of subjective values, especially in contrast to hedonic values, see my reply to Kauppinen, below.)

But DHS claim that, even if the agent doesn't know what the possible lived experience will be like, she can still rationally estimate its value. How can this be true? If the agent is not antecedently in a position to know what the experience would be like, she cannot imaginatively grasp or otherwise represent what it would be like for her to have that lived experience, and so she cannot grasp the outcome's subjective value. This follows from the fact that she cannot grasp and represent the propositions concerning the nature and character of the lived experience.

What's gone wrong here? The problem is that DHS have subtly conflated *intrinsic* value, which they define as value that is had by an outcome "in virtue of its [that outcome's] intrinsic properties" with *subjective* value. And in fact, in the rest of their discussion, their arguments are entirely focused on how one can estimate the intrinsic value of an outcome (for example, by using testimonial and behavioral evidence). Intrinsic values, as they define them, can be communicated by testimony, description, and behavior: no experience needed.

But whether the agent can grasp the *intrinsic value* an outcome might have, a value it has solely in virtue of its intrinsic properties, is irrelevant to my argument that an agent cannot rationally determine her expected subjective value for that outcome.³⁴ Premise 3 is formulated as a claim about *subjective* value. Indeed, all of my arguments concerning transformative experience and the rationality of choice and decision-making are focused on the agent's decisions as decisions framed in terms of the subjective value of outcomes and the expected subjective value of acts. This does not imply that intrinsic values aren't to be included in a global assessment of expected utility. It's just not the type of value that matters most in this context, because it's not the type of value involved in the contents that we are seeking to know when making this life-making decision.³⁵

³⁴ In any case, we need to stay away from affectless, experience-free characterizations of value.

³⁵ See [Moss Unpublished](#) for related discussion, and an argument that transformative decisions can fail a probabilistic knowledge norm.

Indeed, one way to see that subjective value is importantly different from intrinsic value is to draw out how, for many life-changing decisions, extrinsic properties must be part of what is assessed. For example, it doesn't seem to be intrinsically more valuable to be a person who can see than to be a person who is blind. But one might argue that, given the way our society privileges those who can see, it is *extrinsically* disvaluable to be blind. If so, the extrinsic disvalue will likely be determined by the nature of one's lived experience as a blind person in our society, and should be assessed by the agent when she makes her global determination of the subjective value of the lived experience.³⁶

So, one reason why the phenomenon of epistemically transformative experiences does special work is that it causes an agent's subjective value function to be partial. This also helps to clarify why attempting to represent transformative decision-making in terms of imprecise credences is insufficient: the problem, at least in the first instance, is with subjective *values*, not credences. (That said, DHS are entirely correct in their point that the sparseness and messiness of the evidence also leads to problems with credences. It's just that this isn't the only problem, or even the most destructive problem, for transformative decision-making.)

But there is a second important role for epistemically transformative experience: its role in the problem of transformative decisions involving life-changing experiences and the incommensurable preferences they create. The epistemic inaccessibility of the nature and character of her future lived experiences means that the agent facing a transformative decision is epistemically isolated from her future. She cannot look forward and see what it will be like to live in her future circumstances, or indeed, to see who she'll become. This means, first, she cannot neutrally examine her future, incommensurable preferences and compare them to her current preferences in order to develop higher-order preferences for how she wants her life to evolve. And second, because she cannot first-personally know who she is making herself into, she faces a distinctively existential problem, a kind of crisis of rational self-control and self-development.

As DHS point out, "if future preferences are rationally significant for present choices, then this means that one would have to either concede that decision theory is not fully comprehensive as a theory of practical reason or to find a way to extend decision theory so it provides guidance about how to act in light of preference-shift" (319).

Grant that future preferences are indeed rationally and personally significant. Then, in cases of transformative decisions involving life-changing experiences, an agent's current preferences might shift so that she finds herself in new scenarios with new, incommensurable preferences. But, when she faces this decision, the epistemically transformative nature of the

³⁶ For related discussion about the role of social conditions in transformation and the nature of lived experience, see [Barnes 2015a](#).

life-changing experience blocks her from using her first-personal epistemic capacities and knowledge to mentally evolve herself forward and assess these new counterfactual scenarios from her first-personal perspective before she acts. So she lacks the capacity to develop a decision-theoretic guide for how to act.

As I described above, in [section 5](#), the epistemic inaccessibility of the future possibility of radical self-change creates a first-personal, existential analogue of the familiar theoretical issues involving scientific discovery and conceptual revolution. We can describe the problem as a first-personal version of a Kuhnian paradigm shift, that is, where the agent undergoes a first-personal conceptual and preferential paradigm shift as the result of a transformative experience. The combination of epistemic and personal transformation leads to disaster for decision theory as a guide for action. As Bas van Fraassen (1999) puts it:

What is the theoretical problem here for decision theory? Imagine that we contemplate a decision in favor of a certain option, of which the outcome is, by our present lights, ourselves speaking and thinking nonsense, while faring materially much better. Does that make sense? It seems that this would be a true abdication of reason, and not just because we classify it as a bad outcome, an outcome with a low value. Rather, we must doubt that we are coherently framing a decision for ourselves here. For how can we tell that what we *now* see as material welfare in that future will be cognized as such *then*? And if it is not, what about a future in which we are by our *present* lights well off, and by our lights *then* miserable or suffering a great loss? . . . Turn back now to the person totally inside a certain scientific world picture which is becoming burdened with more and more blatant anomalies, severe calculational difficulties, failing predictions, epicycle-laden explanations, and so forth. An alternative appears, some people are beginning to talk about a strange new theory which makes absolutely no sense, and violates the most basic commonsensical expectations of what nature can be like. What is still classified as a satisfactory outcome? To solve the problems of course; taking some absurdity seriously certainly does not count as a solution, and even if it did one would have to be an imbecile to expect it to be vindicated by future experiments. If that person stops a moment to envisage himself converted to the strange new ideas, he sees himself in imagination stooping to irrationalism, he hears himself babbling with (*c'est le bouquet!*) an air of having explained the inexplicable. (75)

The conceptual problem for revolution in theory undermines the idea that the individual proceeds rationally as she makes her first-personal life-changing decisions just as much as it undermines the idea that science proceeds rationally through theoretical and conceptual revolution. Because of the epistemic inaccessibility of the nature of her future, an agent cannot step back and neutrally or selflessly compare her current epistemic situation and her current preferences to her future epistemic situation and her future preferences. She cannot step outside of her conceptual framework or her “preferential” framework in order to develop a consistent guide for the radical shift in perspective and preference that transformative experience implies. Echoing a phrase of van Fraassen’s, there is no first-personal view of the self that is invariant under such transformations.

In closing, while I’ve been critical of DHS’s argument, I want to emphasize that their objections are well-taken, and there is much in their discussion that I find suggestive and interesting. In particular, their discussion in section 3 of their paper raises a new and interesting puzzle concerning the rationality of decisions that might have temporally distant transformative implications.³⁷ They also raise important questions in section 4.1 of their paper about when it can be authentic to use the testimony of experts in place of one’s own first-personal judgments for life-defining choices such as becoming a parent. These are subtle issues about what authenticity can license that deserve further exploration.³⁸

6.2 Social Choice, Social Justice, and Social Identity

Rachael Briggs (2015) explores ways of making sense of wellbeing in contexts of personal transformation. She is interested in working out the structure of intrapersonal comparisons in cases of transformative choice, where the decision is whether to undergo a personally transformative experience.

While the puzzles involved in interpersonal comparisons have been well explored in social choice theory, the problems with intrapersonal comparisons have been neglected, and she sheds new light on the nature of intrapersonal comparisons by assessing them against interpersonal theories of comparisons between different individuals. Her paper gives us an excellent assessment of the terrain for interpersonal comparisons of utility, and gives us an important new way to think about these problems by showing us how to use this map to explore related problems with intrapersonal utility changes, including places where interpersonal utility comparisons and intrapersonal utility comparisons diverge.

In her paper, Briggs sets aside the complication with epistemically transformative experiences in order to focus on the distinctive issues concerning

³⁷ Moss (Unpublished) argues for a solution to this problem.

³⁸ For related discussion about authenticity and expertise in transformative decisions, see Collins 2015, Campbell 2015, Paul 2015a, and Paul Unpublished.

personally transformative preference change. She explores the questions around how we should understand wellbeing in cases of transformative personal experience and transformative personal choices, under the assumption that we have full information about the nature of these transformative preference changes across worlds and times.

Briggs points out that intrapersonal utility comparisons involve the notion of prudence, where prudent choices are those that the agent believes will maximize her overall wellbeing, rather than her wellbeing at just one (short segment of) time. So how we decide to manage intrapersonal utility comparisons is intimately related to how we think about prudence. We might even think of the central complication in personally transformative choice, the problem of how my current self is to regard decisions about what future self to become, as a distinctive kind of prudential complication. Should I remain the self that I am? Or should I become a new self? This raises a further question: how do our decisions about prudential wellbeing map onto first-personal decisions at a time, where those decisions affect our preferences for self-change?

Such questions are versions of the question of whether I should act conservatively or liberally with respect to allowing myself to change. As Briggs points out, “without intrapersonal utility comparisons, preference satisfaction theories are ill-placed to explain why and how I should defer to my future preferences” (202). And once we see the need for intrapersonal utility comparisons, questions arise about how to adjudicate the preferences of my current self versus my merely possible future selves, especially if I currently find those possible future selves repugnant or alien. As Briggs argues, unless our concept of rationality is disappointingly thin, we must find a way to take these future preferences into account when we make decisions. But the question of transformative preference change across selves is exactly this: *how* are we do so, and what is the role of authenticity, intrinsic subjective value, and self-knowledge in all of this?

The issue arises in Briggs’s discussion of the rigidification strategy for intrapersonal utility comparisons across worlds in cases of transformative choice, and for intrapersonal utility comparisons across times in cases of personally transformative change (= personally transformative experience independent of whether the change was chosen).

The rigidification strategy is a strategy that selects a particular preference ordering as privileged. Briggs explores a hybrid strategy as a way to manage cases of personally transformative choice and personally transformative change. If we use this strategy, in cases of personally transformative *choice*, we should rigidify to the actual world, that is, we should assess the choice in light of the person’s actual preferences, rather than her counterfactual preferences. This allows us to say that what is good for the parent is that she has a child—because she actually has a child. Similarly, what is good for the Deaf adult is that he does not have cochlear implants—because he never received cochlear implants. The idea here is that these agents are

better off in their actual scenarios rather than their counterfactual scenarios, because what is good for them is based on their actual preferences. Their counterfactual preferences are irrelevant.³⁹

In cases of personally transformative *change*, on the other hand, we rigidify eternally, that is, we rigidify across times instead of rigidifying to the now. So I'm describing the strategy as "hybrid" because we privilege our local world, but we don't privilege our local time, in the sense that my actual preferences are privileged, but my preferences now are not privileged. "[W]e interpret my preferences as preferences about what happens at *t*, so that what is good for me is to get peanut butter when I have peanut butter cravings, and Vegemite when I have Vegemite cravings" (214).

As Briggs notes, the hybrid strategy has mixed results for an assessment of wellbeing. "Rigidifying allows the preference satisfaction theorist to give intuitively correct answers in a variety of cases involving transformative choice. In cases involving transformative experience [personally transformative change], the rigidifying strategy is less promising, since it requires us to arbitrarily favor an agent's preferences at one time over her preferences at all other times" (p. 213).⁴⁰ The point here is that there is a natural sense in which an assessment of a person's wellbeing can depend on how they are in the actual world but might be independent of how they might have been. But we don't have a non-arbitrary way to say that a person's wellbeing depends on how they are at one time rather than how they are at another time, and in contexts of transformative change we are presented with the need to perform precisely this sort of assessment of wellbeing.

I found Briggs's discussion very illuminating: it highlights the deeper structure of the way we must think about intrapersonal comparisons and raises a cluster of interesting questions. However, I have concerns about the rigidification strategy for transformative choice. There is a natural sense in which we want to privilege the actual world in some cases. Recall the discussion from [section 5](#), of the childfree person who, *ex ante*, finds the possibility of being a happy parent repulsive. In such a case the rigidifying strategy yields the answer that it's best for her to be as she is, even if the scientific evidence and testimony tell her she'd be a happy parent.

But why accept the Panglossian premise that my actual situation *always* has an advantage over my counterfactual situation? Briggs does not defend the rigidification strategy for all cases of transformative choice, she is merely concerned with showing that in some cases, it provides an intuitively satisfying answer. I agree that in the particular cases she describes the answer seems satisfying. But I am unsatisfied with the justification for the rigidification to the actual world, because I don't see how to apply it in a principled way.

³⁹ For an interesting, related argument, see [Barnes 2015b](#).

⁴⁰ In correspondence, Briggs notes that she doesn't endorse the mixed strategy, although she thinks it has a lot going for it.

The question is particularly pressing in the context of the solution to temporal variation in cases of personally transformative change. Consider the case of choosing to have a child. I think Briggs would agree with me that it is parochial to think that satisfying my preference for being childfree at the time when I prefer being childfree is somehow better than satisfying my preference for being a parent when, at a later time, I prefer being a parent. But then why should I privilege my actual, childfree preferences over the preferences of a counterfactual, pro-parental self? Isn't it equally parochial to think that satisfying my preference for being childfree in a world where I prefer being childfree is somehow better than satisfying my preference for being a parent in a world where I prefer being a parent? If so, the rigidification strategy is the wrong one to pursue in this case.

What settles which preferences we should privilege? Which preferences matter most? And why should actuality play a special role in the privileging? My conclusion, drawn from Briggs's rich and interesting discussion, is that in transformative contexts there may be no non-arbitrary way for a preference satisfaction theorist to adjudicate between sets of preferences to give an account of wellbeing and prudence.

Elizabeth Barnes (2015a) develops the idea that whether an experience is transformative can depend on social conditions. She argues that we can distinguish degrees of transformation, and that social conditions can affect one's transformation in a way that has implications for social justice.

Barnes grants that whether a person undergoes a transformation when she has an experience depends on her previous experiences and on what she is like, but she argues that it also depends on the nature of the environment that she is in. Features of an agent's environment can contribute to her transformation because these features can be causally relevant to her psychological response.

I agree with much of Barnes's excellent paper: I will confine myself to drawing out three implications of her central points.

First, Barnes's examples involving disability show how properties of the environment can affect the nature of one's lived experience in deep and far-reaching ways. Consider someone who, as a fully grown adult, is under four feet tall. Given the "standard" sizing of everything from cabinet height to airplane seats to steps, she will stand in relations to her environment that will have a negative effect on her lived experience.⁴¹ This brings out an especially clear way in which the subjective value of an outcome for a person depends on much more than her internal, intrinsic properties. Subjective values concern lived experience of outcomes, or *what it's like to live in this*, understood as an assessment of what it will be like for a person (or for others affected by her decision) to live in the circumstances of some possible outcome.⁴² It can be a matter of social justice to improve

⁴¹ For related points see [Barnes Forthcoming](#).

⁴² Also see [Campbell 2015](#) and [Paul 2015a](#).

the subjective values of lived experience, and by extension, to construct conditions that facilitate certain types of transformative experiences.

Second, the connection to social conditions and social justice ties questions about transformative experience and decision-making to interesting issues in law and public policy.

In many legal, behavioral economic, and social policy contexts, value must be assessed in terms of monetary costs and benefits. Serious problems arise when the values of some outcomes cannot be quantified, that is, when these outcomes cannot be assigned a value.⁴³ If there is no way to assess the utility of those outcomes, then there is no way to calculate an adequate monetary representation of the value (or a monetary quasi-equivalent of that value) that can be used for the relevant cost/benefit analysis. But as a practical matter of fact, in policy contexts, monetary representations of the outcomes are essential in order to be able to assess, manage, understand, and regulate the scenarios they concern.

Examples of nonquantifiability in contexts involving radical transformation abound. Consider values for lived experience outcomes concerning veterans returning from war trauma, or for victims of terrorism, financial collapse, or significant personal injury, or for individuals making decisions in medical and legal cases. If a subjective value cannot be assigned to the outcome in any straightforward way, unless we are to pretend the problem doesn't exist (and thus ignore the needs of citizens we have a responsibility towards) we need some way to construct a proxy value. This is a matter of social justice.

Recent work by Cass Sunstein is intended to address this problem. He identifies the problem and develops the foundations for what he describes as "breakeven" analysis.⁴⁴ Breakeven analysis is designed to show how to partially manage decision models for cases where the subjective values of the outcomes cannot be determined.

The problem of nonquantifiability is a recurrent one in both public policy and ordinary life. Much of the time, we cannot quantify the benefits of potential courses of action, or the costs, or both, and we must nonetheless decide whether and how to proceed. Under existing Executive Orders, agencies are generally required to quantify both benefits and costs, and (to the extent permitted by law) to show that the former justify the latter. But agencies are also permitted to consider apparently nonquantifiable factors, such as human dignity and fairness, and also to consider factors that are not quantifiable because of the limits of

⁴³ I am indebted to Cass Sunstein for discussion of these ideas.

⁴⁴ A related paper (Vermeule 2013) discusses how problems like this can arise in legal contexts. The paper discusses the way judges and regulatory bodies lack appropriate models for recognizing and regulating the problem, and argues that a more rational decision process would recognize the knowledge gaps.

existing knowledge. When quantification is impossible, agencies should engage in “breakeven analysis,” by which they explore how high the nonquantifiable benefits would have to be in order for the benefits to justify the costs. (Sunstein 2014, 1369)

Social and legal theory needs to explicitly recognize the possibility of gaps in value assignments for these types of outcomes in order to adequately diagnose problems, correctly interpret cases, and design models that can manage the unknowns as effectively as possible. Extending Barnes’s point about social justice, recognizing the way that transformative experience can be the source of nonquantifiability and understanding the structure of the epistemic and personal change involved in transformative experience is essential for the proper identification and management of these social conditions, and seems to be part of what’s needed for a just society.

Finally, the importance of properties of the environment to the nature of an individual’s transformative experience highlights the complexity of using empirical data to predict how an individual will respond to a given experience, for the prediction must take the individual’s environment into account. How an individual responds to a transformative experience may depend as much on her environment, and on the particular combination of her physical properties with her micro-environment, as it does on her intrinsic psychological states.

The importance of environment and context to our ability to make accurate predictions for an individual must not be underestimated. This issue arises even when our data is drawn from research that meets the highest standard, for example, when we are doing evidence-based medicine, where many predictions are made on the basis of randomized clinical trials. In a real-life environment the accuracy of a prediction for an individual can be significantly affected by the shift from controlled experimental contexts to messy, real-life situations. Well tested, highly verified models that work beautifully in controlled settings can crash dramatically in real-life contexts, often because the properties of a particular environment affect individual responses in ways the models are not able to predict (Cartwright 2011).

Rachel McKinnon (2015) argues that the decision to change one’s gender is a transformative decision, since transitioning from one gender to another bears all the hallmarks of other types of transformative experiences. I agree that transitioning from one gender to another is a transformative experience, and that the decision to transition is a transformative decision. If so, then before you transition, you cannot know what it will be like to have your new gender, and so you cannot assign a value to your future lived experience with that gender. Moreover, you are constructing a new self with your choice, which will change your core personal preferences. As a result, you cannot rationally choose to transition, if your choice is based on what it will be like for you to have transitioned.

Does this imply that we cannot make such decisions rationally? No. As McKinnon argues, you might find the status quo intolerable. That is, you might find life with your current gender to have a high negative subjective value, and as a result you place a high value on change. As I'd describe the situation, you disprefer the status quo, and as a result you prefer to change your preferences.

McKinnon argues that we need a model for rationally choosing to transition.⁴⁵ I agree. In [Paul 2014](#), I argue that one way to make a rational choice to have a transformative experience is to prefer to discover the new preferences you'll form as the result of having that experience. I describe this as choosing revelation, that is, through action, you choose to reveal to yourself who you become. This, then, is one model for rational, transformative choice: just as you might rationally choose to have a child based on the preference to discover a new self, yourself as a parent, you might rationally choose to transition based on the preference to discover yourself with a new gender.

McKinnon also shows how epistemically transformative experience connects with the literature on feminist standpoint epistemology and situated knowledge. Understanding the nature of epistemically transformative experience connects to issues concerning epistemic trust and epistemic humility in work on oppression and intersectional identity, with further applications in political theory.⁴⁶

Ryan Kemp (2015) explores the rationality of radical personal transformation in the context of debates over the rationality of moral norms and moral self-transformation. As he points out, these debates face a version of the problem with transformative decision-making: how can a person rationally choose to transform her current self into a radically new self?

I find Kemp's discussion interesting and provocative, and was particularly interested in the connections he draws between the work of Sartre and Kierkegaard and contemporary philosophical issues concerning self-transformation. It may well be true that radical self transformation derives from contingent facts and that it involves a leap of faith. As [Barnes \(2015a\)](#) points out, such transformation may also depend on properties of the social environment, and as [McKinnon \(2015\)](#) argues, the individual may feel that she has no choice but to undergo transformation.

I was unconvinced, however, by Kemp's argument that radical self-transformation cannot be covered by a choice based on revelation, that is, by a choice to transform oneself based on one's preference to discover a new self. Kemp argues that cases of radical self-transformation are cases where

⁴⁵ While I don't provide a model in [Paul 2015b](#), I do develop a model in [Paul 2014](#). The dates notwithstanding, [Paul 2014](#) was not published when McKinnon's article was accepted for publication while [Paul 2015b](#) was widely available online from January 2013. Thus McKinnon is justified in not engaging with the (2014) book.

⁴⁶ For discussion of the problems that transformative experience raises for accounts of democratic ideals that rely on cognitive empathy, see [Stanley 2015](#), 102–105.

a person explicitly decides to make a change with the precise intent of uprooting her central preferences. To put the point a bit more colorfully, transformative experiences involve a decision to risk normative death in order to experience something new; radical self-transformation involves a decision to embrace normative death at the outset. (Kemp 2015, 395)

If I understand Kemp correctly, he wants to distinguish between transformative decisions where a person makes a decision under conditions of uncertainty as to whether a consequence of her act is self-replacement, and transformative decisions where a person is certain that a consequence of her act is self-replacement. This is an interesting distinction that may well have normative implications in the moral domain. However, one can rationally choose self-replacement based on a preference to replace one's self with a new self in both types of conditions. If the choice to φ is transformative, one can rationally choose to φ if one's choice is made based on the higher-order preference to discover what it's like to become a new self. Such an act meets the normative standard for rationality even if it creates problems for other types of norms.

Muhammad Velji (2015) discusses the question of accommodation for those with religious preferences, such as accommodation of those who prefer to veil or accommodation of those who prefer to keep kosher. He argues that we should not refuse to accommodate religious preferences on the grounds that such preferences result from the choice to become religious. His interesting argument turns on the sort of epistemic and personal transformation involved in the slow process of religious transformation through pious engagement and practice.

Velji argues that we should not see the choice to train oneself in religious piety as a fully informed choice, for the self-development involved in the transformation of oneself into a pious believer can change a person into a new self with new preferences, but the preferences of this future self are epistemically inaccessible before her religious transformation is undertaken. If the prospectively religious believer cannot know whether she will require religious accommodation, such as accommodation for veiling, until she achieves a certain level of piety, and once she reaches that level of piety, veiling is part of her religious identity, then her choice to believe is not informed in a way that undermines her right to religious accommodation.

The central idea is that, while the choice to be pious is indeed a choice, the choice to pursue religious piety should not be regarded as a choice analogous to choosing a particular lifestyle such as choosing to drive a fancy car or choosing to develop one's physical prowess through skiing. Rather, it's a choice that can transform a person. If so, veiling and other religious practices should be accommodated, for they are practices that are constitutive of one's self-identity, not mere lifestyle choices, and the

nature of one's religious self-identity (especially, the nature and extent of the accommodation needed) is only discovered after the path to piety has been embarked upon.

6.3 Subjective Value and Happiness

In his engaging paper, Antti Kauppinen (2015) argues that (i) “nonexperiential” values are far more important for our big life choices than “experiential” values, and (ii) choice based on valuing revelation is not normatively acceptable.

Kauppinen's primary target is my notion of subjective value. As I define subjective value, it is the value of lived experience. This fact seems to have been misunderstood by some of my critics, so I welcome the opportunity to clarify the idea. In particular, Kauppinen, along with some of the other ethics-oriented papers in this volume, such as those by Chang, Dougherty, Horowitz, and Sliwa, and Harman, take subjective value to be something it is not. Kauppinen thinks it is the value of mere subjective feel, and rejects it as a suitable value for making life choices. But lived experience is much more than mere subjective feel.⁴⁷

While the idea that we value lived experience is very natural and intuitive, it *is* unfamiliar from the perspective of contemporary practical ethics, especially since much of “analytic” philosophy, with the exception of the philosophy of mind, has traditionally regarded phenomenology, and talk of experience more generally, with fear and trembling. Moreover, my arguments for the necessity of experience in generating the epistemic capacity to imagine new types of future lived experiences exploit classic examples from the philosophy of mind, but those examples were traditionally used in arguments concerning mere qualitative feel. So it might seem that I am arguing that big life choices should be made based merely on what sort of phenomenology they create, and not on more important bases, such as the objective value of those outcomes. This misrepresents my project in a significant and far-reaching way.

Clarification is in order. As I'll explain below, my argument is that experience is necessary for our ability to represent and imagine the nature and character of outcomes that involve much more than mere subjective feel.⁴⁸ Such outcomes include those involving love, betrayal, fear, friendship, loyalty, personal sacrifice, etc., or what I describe as “rich, complex, experience-involving” states of the world, and are the sorts of outcomes that Kauppinen and others are very rightly concerned with.

Let's start with whether the concern is merely about our inability to imagine mere subjective feels. As John Campbell (2015) points out:

⁴⁷ See Campbell 2015 and Paul 2015a for further discussion.

⁴⁸ Barring special machines that can create brain states that would duplicate the brain states we'd get from having the experience, etc.

Although current philosophy of mind has long recognized the importance of imaginative understanding, it's been given a remarkably restricted role: providing one with knowledge of the qualia, the purely internal characteristics of someone's mental life.

But there is no need to restrict imaginative understanding to such a confined role, and my arguments do not do so.⁴⁹ Instead, I emphasize the central importance of imaginative representation in prospective planning, anticipation, and decision-making, and show how enlarging the role for imagination enlarges both the interest and the scope of transformative decision-making.

That is, in many situations, including situations where we are struggling to assess possible life choices, we want to be able to imaginatively represent new types of outcomes in which we experience new events and live in the world in new ways, including ways we'd experience ourselves in such outcomes. My argument is that well-known arguments from the philosophy of mind about the need for experience to represent possible qualitative states should be extended to our ability to represent these new outcomes. In particular, the role of experience is just as important in imaginatively representing the nature of possible states in which we experience and live in the world in new ways as it is when imaginatively representing fairly simple qualitative states. And, importantly, possible states in which we experience new events and new ways of living in the world include some of the rich, complex, experience-involving states of the world that concern us most when we make big life decisions.

Why is experience necessary for a grasp on the nature of these rich, complex, experience-involving states? As follows: To grasp the (salient) nature of a complex, many-featured state of the world, one must grasp the nature of the (salient) features that compose or constitute this state.⁵⁰ In the complex states of interest, such as outcomes that involve love, betrayal, friendship, aesthetic beauty, sacrifice, etc., some of the salient features that partly compose or constitute the states are fundamentally experiential: they are sensory, such as what it's like to feel pain, or they concern internal experience, such as what it's like to feel certain basic emotions or what it's like to hear beautiful music, or they include our experience of our more complex physiological and psychological responses to various events or properties in the world.

Other salient features that may partly compose or constitute the states are not primarily experiential, but our ability to grasp their natures depends on our experience, such as our ability to grasp what it's like for another

⁴⁹ For further discussion see [Paul 2015a](#).

⁵⁰ This is a necessary condition. Sometimes we can have a grasp on the natures of the parts, yet the nature of the whole continues to elude us. Also, I'm using the term "salient" to gloss the possibility that there might be features that make a largely irrelevant contribution to the nature of the state. These are not features we need to be concerned with.

person to feel pain or to see color. (For example, we have to know what it's like to feel pain to empathize with the pain of another.)

Of course, there may be other features that partly compose or constitute these complex states that are not experiential, nor is our grasp of their natures dependent upon our grasp of something experiential. We can call such features *wholly non-experiential*. But unless the salient nature of a state is wholly composed of wholly non-experiential features, to fully grasp its salient nature, we will need experience.

And finally, the states we care about when we make big decisions, such as states of the world that concern the fear, anger, friendship, love, aesthetic beauty, sacrifice, etc., of ourselves and others, are not wholly composed of wholly non-experiential features. They are not “experience-free” states, and their salient natures are grounded, at least partly, by the natures of their experiential features. Love and friendship, for example, involve experiences and feelings towards others and responses to others, as does the attachment between parent and child. Most of the complex states involving new events and ways of living in the world that we want to assess in life-changing situations are constructed from all three types of features, and in most, experience is highly salient.

Now we can see the role of the argument from experience: the rich, complex, and meaningful states of the world that concern us in many big life decisions are not experience-free states, they are experience-involving states. In order to imaginatively represent the nature of these experience-involving states, it is necessary for a person to have had the right type of experience. Without having had the right sort of experience she cannot represent the nature of some of the most important features that compose the state, and thus she cannot represent the (salient) nature of the state itself.

So the idea is that experience of the relevant kind is needed for one to have the epistemic capacity to grasp the nature of the experience-involving states involved in big life changes, for it gives us the imaginative capacity to first-personally represent or model our possible future selves (and the possible future selves of others) in those states. And note that subjective value is intended to be understood *de re*, in the sense that they concern what it's like (for oneself and for others) to live in these possible circumstances, or as I sometimes describe it, “what it's like to live an outcome” in the world.

As I indicated above, I connect the nature of these experience-involving states with subjective value, describing it as “the value of lived experience” or the value of what it's like to live in possible outcomes. The idea is that only by grasping the nature of experience-involving states do we have the capacity to represent and assign them subjective value.

This is why arguments against subjective value as grounded solely by mere subjective feel miss the mark. Arguments based on “experience machine” worries that suggest that experience isn't important are also misguided. My argument is, first, that subjective values of the important types

of lived experiences, understood *de re*, are among the values that matter to us, especially when making life-defining transformative choices. Second, that experience is *necessary* to grasp these important subjective values. I am not arguing that experience is sufficient to grasp these subjective values, nor that some sort of thin, purely internal, affectless phenomenal character is sufficient to ground subjective value.

You might think that objective value is ultimately what we should care about. I do not object to this, given that, for experience-involving states, their objective value often depends in substantive ways on their subjective value. (See my discussion of Ruth Chang's paper in [section 6.4.](#)) One might also wish to assess the intrinsic value of these states, as [Harman \(2015\)](#) and [Dougherty et al. \(2015\)](#) do. Again, I have no objection to this, as long as we are clear that, somehow, when assigning objective values to outcomes, we will need to assign them subjective values. This is because in transformative decision contexts, subjective value is of central importance to the decision maker.

Kauppinen's discussion might seem somewhat orthogonal to all this. He says that he is concerned to assign values that he describes as "non-experiential." But Kauppinen also wants to assign values to states that include "integrity, commitment, friendship, meaning, or achievement," just as I do. Where is the significant difference? Mostly, I suspect, in how we are drawing distinctions. In particular, Kauppinen argues that experiential value is merely (prudential) hedonic value, and as a result thinks it is of little importance. But since my experience-based subjective values are not merely hedonic values, I am inclined to value many of the same features of the world as he is.

For these reasons, I agree with Kauppinen that prudential goods like achievement, friendship, and self-respect are the sorts of things we should assign value to when we assess the expected value of acts involved in major life decisions. It's just that valuing these goods for transformative decision-making involves assigning them subjective value, and that's the type of value we need to focus on in this context.

Attempting to value outcomes involving friendship, love, and self-respect without including lived experiences in these states leads to a weird sort of zombification of what we are supposed to value.⁵¹ The zombie equivalent of love is all right, I suppose, but it's not the kind of thing that's suitable for prudential value. We want to value real love, and real love involves conscious experiences and emotional relationships and many other sorts of things that subjective value is designed to capture. (Again, this does not mean that love is merely experiential or is merely a subjective feel.)

In correspondence, Kauppinen grants that achievement, friendship, and the like necessarily involve experiences, but maintains that an important

⁵¹ Mark Johnston (2006) makes related criticisms of what he terms the "Wallpaper View," where sensation is treated as a mere add-on to perceptual judgment.

part of their value is grounded in their non-experiential features. On this much, we agree. Where we disagree concerns what we have to know or be able to represent. On my view, the subjective values of lived experience outcomes are among the values that matter most in these decision contexts, and first-personally grasping such values requires experience-based knowledge, or at least, an experience-generated epistemic capacity. On Kauppinen's view, experience-based knowledge or capacities are not required for grasping the values of lived experience outcomes, that is, he thinks experience is not needed to first-personally grasp the aspects of lived experience outcomes that are relevant to their value. What our debate brings out is a deep difference in our views about the epistemic and conceptual requirements for assigning value to outcomes involving lived experience, and raises further interesting questions about how to understand the metaphysical structure of lived experiences and the grounding of the relevant values.

Kauppinen also criticizes the adoption of revelation, or discovery, as a subjective value. I see his concern, and indeed, I do not think that choosing based on revelation alone is an especially successful way to resolve the problems with transformative decision-making. Simply choosing a particular life path to discover what that path will be like is not at all how we normally want to proceed, especially given the modern focus on planning one's future and thinking in terms of narrative goals. So I agree with Kauppinen that it can be normatively unacceptable to choose a life path for revelatory reasons alone. But the norm that's violated is not an epistemic or a rational norm, at least, not as long as you are choosing consistent with your rational preferences. Rather, it's a norm of authenticity, because we normally prefer to know the nature of the life path, at least to some significant degree, that we are choosing. And the main problem with transformative decision-making, as I've been at pains to point out, is that the norms of authenticity conflict with the norms of rationality, and there is no easy or obvious way to resolve this conflict.

6.4 Decision-making in Contemporary Ethics

Elizabeth Harman (2015) explores the relationship of transformative choice to her groundbreaking work on "I'll be glad I did it" reasoning. She discusses two cases that we both find to be of great interest, the case of a woman choosing to have a child and the case of a parent choosing to implant their deaf child with a cochlear implant. Harman argues that experience is not necessary for people to rationally make decisions in cases like these, and argues that reliance on reliable testimony, as opposed to reliance on faulty "I'll be glad I did it" reasoning, will allow us to make rational decisions in cases of transformative choice.

In reply, I argue that experience of the right sort is necessary if we choose based on our assessments of the expected subjective value of becoming a parent. I will discuss the structure of "I'll be glad I did it" reasoning and

show how the structure of transformative decision-making and the values it concerns are importantly different from those involved in faulty “I’ll be glad I did it” reasoning. (Exposing the difference between these structures is also important for my discussion of [Howard 2015](#), below.)

After I distinguish the structure of transformative decision-making from “I’ll be glad I did it” reasoning, I’ll explain why testimony can’t provide the relevant information to the agent who lacks a subjective value function. I’ll close by exposing how an ambiguity in Harman’s claim that it is “better” for a deaf child to have a cochlear implant illustrates the way that deep questions concerning this sort of real-life choice concern epistemic and personal transformation and the value of subjective lived experience, not merely testimony and faulty “I’ll be glad I did it” reasoning.

To set the stage for her defense of the view that people can rationally make transformative choices by relying on testimony, Harman argues that becoming a parent is not epistemically transformative. Her argument is that (i) she had parenting-like experiences with her much younger sister and the child of a close friend, so she knew what it would be like to become a parent before she had a child, and (ii) she knew before she became a parent that she would experience joy and other sorts of emotions when she had her child.

Harman’s personal experience does not refute the argument that one has to stand in a parent-child attachment relation to a child to know what that distinctive type of experience (parenting experience) is like. Why? Because, by her own admission, Harman had the experience of being a parent before she physically produced her own child, through experiences with her younger sister and with her friend’s baby. She notes that she had feelings for her sister that were “parental in their nature” and when describing her feelings for her friend’s baby says “I experienced my own love for her baby, which was unlike any feelings I had ever had (as an adult) for a baby” (326). If we take this description of her experiences at face value, Harman *allopanted* her friend’s child (and her younger sister): that is, she loved and cared for children that were not her own.

Alloparents are individuals who are not the biological parents of a child, but who stand in a parent-type relation to that child. So Harman had experience with being a parent, just not of being a parent of a child she’d physically created. But my argument does not require that one must be a biological parent to know what it’s like to be a parent. My argument is just that the type of loving attachment to a child that one experiences as a parent is a distinctive experiential kind. If you’ve had experiences that are instances of that kind, you can know what it’s like. Just like you can know what it’s like to taste Vegemite once you’ve tasted Marmite (both are yeast extract spreads),⁵² if you parent other children before you parent your own, you can know what it’s like to be a parent.

⁵² Well, many Australians and English would deny that Marmite and Vegemite taste anything like the same. And New Zealanders have their own kind of Marmite. I get it. But bear with me for the example: most Americans can’t taste the difference.

While my argument is philosophical, these questions are being explored in the psychological literature as well. Preliminary empirical results from the work of Josiah Nunziato and Fiery Cushman indicate that many people report having transformative experiences when they experience significant life events such as becoming a parent. In particular, people report that the experience changed them in ways that they could not foresee. In addition, further results suggest that people are over-confident about their ability to anticipate the changes that will occur in their beliefs, preferences, desires, and values as a result of a transformative experience.⁵³

There is another problem with Harman's inference from her own case. The experience of being a parent is a very broad type of experience, one that every parent who forms an attachment to her child has acquaintance with. But within the parental experience-type, there may be further distinctions to be made, for there are many distinctive types of parenting experiences: parenting an extremely gifted child, parenting a severely physically disabled child, parenting a terminally ill child, parenting a mentally ill child, etc. Each of these subtypes of parenting experience may be distinctive and different enough from the others to require experience of that particular subtype in order to properly assess their subjective value. (This is a reason why having a second child can be transformative.)

Given her description, Harman was *lucky* with the type of parenting she had experience with: that is, her account suggests that her child was relevantly similar to the children she alloparented, and that she responded to her child in a way that was relevantly similar to the way she'd responded to the other children. But if something had been different, for example, if the child she physically produced had been distinctively different from the other children she'd alloparented (e.g., by being terminally ill), her alloparenting experience might not have been similar enough in the relevant way for her to grasp a significant part of the subjective value of that outcome. This brings out a mistake in her description of an argument she attributes to me.⁵⁴ My argument for the epistemically transformative nature of parenting experience is not based on the claim "that one cannot know ahead of time, regarding any specific possible experience of pregnancy and parenthood that one may have, what *that experience* would be like." Rather, the argument is that one needs experience of the relevant *type* of parenthood experience in order to be able to represent experiences of that type in order to grasp their subjective values.

The second problem with Harman's argument against parenting as an epistemically transformative experience concerns the distinction between knowledge-that and knowledge derived from experience.⁵⁵ There is no

⁵³ I'm indebted to discussion with Fiery Cushman here.

⁵⁴ Harman constructs her own version of the argument, so I am focusing on (and quoting) the version of the view she describes. The arguments I actually develop in [Paul 2014](#) and [Paul 2015b](#) are not quite the same.

⁵⁵ For a related objection to my view, see [Krishnamurthy 2015](#).

question that there is a lot of knowledge that you can have before you decide to have a child, such as knowledge that it will probably be tiring, knowledge that you are likely to love your child, and knowledge that you could have less time for favorite hobbies. This was never in dispute. But Harman seems to think knowing-that it is likely that one will have these experiences, via moral testimony and via observing others, is enough to know what it's like to be a parent. In particular, she seems to think knowing that one is likely to feel the types of emotions that many associate with parenthood is enough to assign *being a parent* subjective value. "There is definitely a kind of joy I had never experienced until now. But I knew there would be" (326).

However, knowing *that* you will (probably) feel new emotions isn't knowing what it's like to feel them, just like knowing that you'll see red isn't knowing what it's like to see red. Her arguments for the role of testimony in gaining knowledge suggests that, like Dougherty et al. (2015), Harman is implicitly replacing assessments of subjective value with other types of value, such as intrinsic value, which is not the sort of value my arguments for choosing to become a parent concern. So her claims about the knowledge that we can have before a transformative experience do not constitute an argument against my view that people cannot rationally choose to become parents based on their assessment of the expected subjective value of that act. Moreover, she makes general, problematic assumptions about what we can know about our own case from knowing anecdotes and statistics about other people.⁵⁶ I will return to the question about the types of value judgments we are making in my discussion of cochlear implants, below.

In the second section of her paper, Harman discusses reasoning involving decision-making.⁵⁷ One part of her discussion concerns a distinctive kind of reasoning, involving "I'll be glad I did it" considerations, where a person reasons from the fact that she is attached to the outcome she actually chose to the claim that she should have chosen that outcome. I agree with Harman that this sort of reasoning is faulty.

But there is a second part of her discussion in section two that deserves a more critical assessment: her discussion of transformative decision-making, where a person makes a decision to undergo an experience that transforms her core personal preferences. Transformative decisions include decisions to become a parent and decisions to radically change one's sensory capacities, such as the decision to get a cochlear implant (if one is congenitally deaf). The cochlear implant case is complicated by the fact that, usually, parents must make the transformative decision for their very young child. In

⁵⁶ See my reply to Dougherty, Horowitz and Sliwa for further discussion about intrinsic versus subjective value. For a case study illustrating the dangers with naïve interpretations of empirical results, see my reply to Sharadin 2015. Finally, for detailed discussion about problematic inferences from anecdotes and statistical results to one's own case see Pettigrew 2015, Paul 2015a, Moss Unpublished, and section 3–section 5 above.

⁵⁷ Thanks are due to Matt Kotzen for very helpful discussion.

Paul 2014, I argue that a central problem for such decisions concerns the incommensurability of the individual's preferences before and after the decision, and the epistemic inaccessibility of the first-personal perspective of the future, transformed self from the current, untransformed self's first-personal perspective. (I discuss the general structure of this problem in section 5, above.)

Harman claims that we can rely on the testimony of satisfied parents or other "experts" as a guide for transformative decision-making. However, as I argue in reply to Dougherty et al. (2015) and others, and as I discuss in section 5, above, the agent cannot use such testimony to know whether she is making a decision that will satisfy her current preferences, or whether it simply satisfies the future preferences of a future, alien self.

Is, as Harman's discussion suggests, the real source of the problem with using testimony merely that agents can be susceptible to the distorted sort of reasoning based on attachment we see in some of the "I'll be glad I did it" examples? Is it correct to think that, once we set this sort of reasoning aside, we can use testimony to resolve cases of transformative choice, such as choosing to become a parent or choosing a cochlear implant?

No. The faulty "I'll be glad I did it" reasoning Harman discusses is not what is creating the deep problem for rational choice in cases of transformative decision-making. What's going on in cases of transformative decision-making (which include real-life cases of transformative decision-making, such as the choice to receive a cochlear implant) is a different problem.

I'll show this by comparing the structure of Harman's cases to the structure of cases of transformative decision-making that create the deep problem for rational choice.

Harman's first case involves an exam. In this case, at $t1$, you decide to skip a film and study for your exam instead. While you might enjoy seeing the film, you have a higher-order preference to skip the film and study instead, a preference to do well on the exam rather than see the film. At $t2$, you are glad you studied at $t1$. You were right to study for the exam at $t1$, because your preferences, including your higher-order preferences, are consistent across $t1$ and $t2$.

Harman's second case involves a woman who had a baby as a teenager. In this case, at $t1$, she chooses to have the baby. At $t2$, the woman is glad she has the baby. Harman says:

A woman who became a parent as a teen might say, truly, "I should not have had a child as a teen. But I love my son and I'm so glad I did, because otherwise I wouldn't have had him. That I love him and am glad to have had him—that I would not wish to change anything for myself—in no way makes me think that teen parenthood is a good choice for anyone to make." (2015, 335–336)

In Harman's second case, the woman's first-order preferences conflict with her higher-order preferences. At $t1$, she chooses to have the baby. At $t2$, she prefers to have the baby because she has formed a reasonable attachment to him. But her higher-order preference to not be a teenaged mother remains consistent across $t1$ and $t2$. (We could also tell a story where her first-order preferences changed from $t1$ to $t2$, perhaps because she had the baby against her will. At $t1$ she did not prefer to have the baby, though at $t2$ she prefers to have the baby. Still, her higher preference is consistent across $t1$ and $t2$.) Because her higher-order preference to not be a teenaged mother remains, it is still correct, at least with respect to this higher-order preference, that she should not have had the baby at $t1$.

What's the structure of transformative decisions such as the choice to give one's congenitally deaf infant a cochlear implant or the choice of a mature adult to have a child? These cases of transformative decisions are importantly different from the exam case and the case of the teenaged mother. In these cases, the agents' preferences, including their high-order preferences, are transformed as the result of the choice. Cases involving transformative decisions involve incommensurable preference changes across time, not higher-order preference change accompanied by cross-temporally consistent higher-order preferences.

Let's consider the choice of an adult, Anne, to have a child. At $t1$, she does not prefer to have a child. Moreover, all of her higher-order preferences are consistent with this: she prefers not to have a child relative to her other preferences, she prefers to prefer not having a child, etc. Yet, the expert tells her that at $t2$, if she has a baby, she'll have preferences that are satisfied then.

Harman is correct to argue that a bad way for Anne to reason about the expert's testimony is to think that the fact that she'll "be glad" to have the baby at $t2$ is a reason she should decide to have a baby at $t1$. For if Anne has a higher-order preference to remain childless that remains consistent across $t1$ and $t2$, her choice violates this preference, despite a higher-order preference to have the child that is created at $t2$ by the existence of the baby.

But Harman suggests that, once we set aside faulty "I'll be glad I did it" reasoning, we can simply accept expert testimony as a guide to action in transformative decision-making.⁵⁸ This is a mistake. Faulty "I'll be glad I did it" reasoning is not what creates the deep problem for transformative decision-making. This is because the source of the problem with faulty "I'll be glad I did it" reasoning, as Harman describes it, involves a distortion due to attachment, with a consequent first order preference change.

⁵⁸ Note that this could require a person to replace her introspective assessments with expert testimony. I discussed problems with this strategy in [section 3](#) and [section 4](#) above. Deciding based solely on expert testimony about one's future preferences would also conflict with the rigidification strategy discussed by [Briggs 2015](#).

Such distortion can occur even when the agent has consistent higher-order preferences across the change, as in the teenaged mother case.

The deep problem for transformative decision-making is different: it comes from transformations that create higher-order preferences that are *inconsistent* with those of the agent prior to the change, and a transformed first-personal perspective that is inaccessible to the agent prior to the change. That is, it involves incommensurable core personal preferences, including higher-order preferences, across the transformative epistemic change of the agent. Testimony won't allow us to evade this problem. (See my discussion of [Howard 2015](#) for a related point.)

Go back to Anne, who is choosing whether to have a child. At $t1$, she does not prefer to have a child. Moreover, all of her higher-order preferences are consistent with this first order preference: she prefers not to have a child relative to her other preferences, she prefers to prefer not having a child, etc.

How should Anne regard the expert's testimony? The trouble is that, in the transformative case, the radical transformation she'll undergo when becoming a parent means that her preferences at $t2$, including her higher-order preferences, will be incommensurable with her preferences at $t1$. So if Anne decides to have the baby, she violates all of her preferences at $t1$. Moreover, her transformed first-personal perspective at $t2$ is inaccessible to her at $t1$. And so the expert testimony that her preferences will be satisfied at $t2$ cannot guide her rational action at $t1$. (A related way to see the difficulty is to explain that Anne cannot assess the relevance of the expert advice to her current self, because the expert advice merely concerns the preferences of her possibly transformed self, a self that Anne at $t1$ regards as psychologically alien.) Adding insult to injury, if Anne decides not have the baby, she violates all of the preferences she'd have had at $t2$ if she had had the baby.

The problem with incommensurable, inaccessible future perspectives also infects the structure of cases involving cochlear implants. Such cases are best thought of as forced choices, because parents must decide whether to implant their child when the child is very young. The comparison of interest is between very different, distant future outcomes: an outcome with a (significantly older) Deaf child, and an outcome with a (significantly older) child who can hear in a species-typical way.

Consider a parent who must decide whether to implant her congenitally deaf infant. Let us assume, first, that the parent does not have a higher-order preference to have a child who can hear, nor does she have a higher-order preference to have a Deaf child.

Assume that the parent chooses to reject the implant at $t1$. At $t2$, the Deaf child is glad to be Deaf. Is it correct to say that the defense of the decision to reject the implant at $t1$ must involve faulty "I'll be glad I did it" reasoning? No. Unlike the teenaged mother, this parent does not have a

higher-order preference across $t1$ and at $t2$ conflicting with her higher-order preference to reject the implant at $t1$.

Now complicate the case. Perhaps, at $t1$, an expert tells the parent that, if she implants her child at $t1$, that at $t2$, the child will be glad to have the implant. But if the parent refuses the implant at $t1$, at $t2$, the child will be glad to be Deaf. What's the right assessment for the parent to make?

The different preferences at $t2$ stem from the fact that receiving a cochlear implant is transformative for the child. If he receives the implant, at $t2$, he'll have one set of preferences, P_H , formed in part by his life as a person who can hear. If he does not receive the implant, at $t2$, he'll have a different set of preferences, P_D , formed in part by his life as a Deaf person. So if the child is implanted, at $t2$, the parent and the child have preferences (P_H), including higher-order preferences, that the child is implanted. If the child is not implanted, at $t2$, the parent and the child have preferences (P_D), including higher-order preferences, that the child is Deaf. Preferences P_H are incommensurable with preferences P_D .

In this case, if there is no preference had by the parent at $t1$ prohibiting implantation, choosing to implant is rationally permissible. Likewise, if there is no preference had by the parent at $t1$ prohibiting the refusal to implant, refusing to implant is rationally permissible.

We can complicate the case further. A parent might have preferences, including a higher-order preference, to refuse the implant. The expert might tell the parent that, were the child to be implanted, at $t2$ the parent and the child would have preferences P_H , including higher-order preferences, that the child is implanted. Does this imply that the parent should choose to implant? No. If the experience of becoming a person who can hear (and of being the parent of that person) is transformative, then the parent's preferences at $t1$ are simply incommensurable with her preferences at $t2$. And as a result of the radical preference change that could occur, the parent at $t1$ can regard the self she'd be at $t2$ as alien to her current self at $t1$. (And she might regard who the child would become at $t2$ as alien to who the child is now, at $t1$.) So testimony will not evade the problem for rational decision-making for a parent faced by this sort of transformative choice.

Is it nevertheless "better," as Harman suggests, for the child to receive the implant than to forgo it? In what sense could it be better? One way it could be better is that the parent's (and child's) preferences could be better satisfied in one outcome than they are in another. But as we have seen, preferences P_H are satisfied in the outcome where the child is implanted, and preferences P_D are satisfied in the outcome where the child is not implanted. Moreover, preferences P_H are incommensurable with preferences P_D . So the sense of "better" with regard to better satisfying one's preferences does not apply.⁵⁹ Is there another sense of "better" in play here? In particular, is Harman suggesting that being able to hear in a species-typical way is

⁵⁹ For related discussion see [Barnes Forthcoming](#) and [Howard 2015](#).

somehow *intrinsically more valuable* than being Deaf? Is this what is supposed to solve the problem for the parent faced with the transformative decision to implant her deaf infant? Her defense of the role of testimony, and her approving citation to [Dougherty et al. 2015](#), suggest this possibility, but she does not explicitly endorse it in this paper.

I categorically reject the thesis that being able to hear in a species-typical way is somehow intrinsically more valuable than being Deaf. If Harman is arguing that, even in a world where cochlear implant technology is perfect, parents should give their deaf children cochlear implants because being able to hear is intrinsically more valuable than being Deaf, we need an explicit and compelling argument for that view.⁶⁰ And once we move to cases in the actual world, what deserves further scrutiny are the possibilities that (a) being Deaf is extrinsically less valuable than being able to hear in a species-typical way, given the way that many societies are organized, and (b) being Deaf is extrinsically more valuable than being implanted, given technological and other facts about cochlear implants.⁶¹ Such questions about extrinsic value (and the role of the subjective value of lived experience in contributing to extrinsic value) are difficult but highly salient in this context.⁶²

In sum: the deep questions surrounding big life decisions and the transformative choices they involve concern epistemic and personal transformation and the subjective value of lived experience. Harman's arguments that we can make these decisions rationally simply by relying on testimony about the intrinsic value of outcomes conflates subjective value with intrinsic value, misdiagnoses the source of the problems, and fails to recognize the deep problems with preference change and alienation when relying on expert testimony to make life-changing transformative decisions.

Dana Howard (2015) discusses the relationship between "I'll be glad I did it" reasoning and transformative decision-making. She argues, contra

⁶⁰ See [Harman 2009](#) for related discussion.

⁶¹ See [Barnes 2015b](#) for related points about transformative experience and social conditions.

⁶² In [Paul 2014](#), I argue that parents without the experience of being Deaf cannot assess the subjective value for their child of being Deaf. I also argue that parents without the experience of species-typical hearing cannot assess the subjective value for their child of being able to hear in a species-typical way. I then argue that, as a result, we should not expect parents to be able to have rationally defensible preferences concerning these subjective values when they make the decision whether to implant. Moreover, in real-life cases, especially given the limitations of current technology, a pressing and important concern for parents with a deaf child involves their future ability to fully communicate and engage with their child, which brings with it the fear of alienation from one's own child. Often, parents prefer to keep their child in the same community as they are, in hopes of maximizing the child's chances of a happy and successful childhood and subsequent preparation for adult life. Thus, Deaf parents may refuse the implant, judging that their child is best off as a member of the Deaf community, and hearing parents may choose to implant, judging that their child is best off as a member of the hearing community. Consideration of facts like these could be part of what leads Deaf parents to form a higher-order preference to have a Deaf child, and for hearing parents to form a higher-order preference to have a hearing child.

Harman (2009), that “the fact that one will be glad one did it never offers up a conclusive reason to believe that one should do it” (358).

In her argument, Howard distinguishes between preference change based on adaptive preference formation, understood here as stemming from a diminished set of life options, and preference change based on other sorts of reasoning. She argues, rightly, that when a disabled person has a preference to be disabled, that we should not assume that that the disabled person’s preferences stem from adaptive reasoning. To show that adaptive reasoning is the source of the preference we must first show that the life options of the disabled person are in fact diminished.⁶³

She then gives a careful and rigorous diagnosis of what goes wrong in paradigmatic “I’ll be glad I did it” reasoning. As she points out, it is morally impermissible for parents choose to disable their children or to permit them to remain disabled based on reasoning about future preferences that stem from adaptive preference formation. She distinguishes such cases from the case of cochlear implantation, holding that the question of what is morally permissible in cases like these is still open. I agree with Howard’s conclusions about moral permissibility and faulty “I’ll be glad I did it” reasoning.

However, as should be clear from my discussion of [Harman 2015](#), above, the basic structure in “I’ll be glad I did it” reasoning can differ dramatically from the structure of a case involving transformative decision-making. As a result, it is not clear how Howard’s conclusions apply to transformative decision-making, especially to the problems concerning the inaccessibility of the subjective value of lived experiences in cases of transformative change. (In [Paul 2014](#), I discuss this problem in the context of decisions involving cochlear implantation.)

Recall that, when making a transformative choice, the act the agent performs can lead to events that transform her self-perspective. This is not a mere change in preferences; it is a transformation of one’s epistemic capacities and a replacement of some core personal preferences. In transformative change, the self that results from the transformative choice can have preferences that are incommensurable with those of the earlier self. Moreover, the preferences of the transformed self can include the preference to be that transformed self.

Although Howard is correct that both “I’ll be glad I did it” reasoning and transformative choice involve changes in a later self caused by an earlier self, the problems of transformative choice are not created by adaptive preference formation.⁶⁴ Rather, the source (and the magnitude) of the problem with transformative decision-making arises from the inability of the self to endorse epistemically inaccessible preference change across

⁶³ See [Barnes Forthcoming](#) for related discussion.

⁶⁴ Problems with transformative decision-making remain even when there is no constriction whatsoever in the life options of the person who undergoes (or who refuses to undergo) the transformative experience.

contexts of radical self-change. Using Howard's notion of "endorsement," the problem with transformative choice is that, unless you can endorse the preferences of your future self *before you change your current self*, your future self's preferences cannot justify your choice at $t1$.⁶⁵ (This is why, contra Howard's assertion [2015, 369], relying on testimony about your future self's preferences isn't an easy way out of the problem of rationally deciding to become a parent. You lack the ability to endorse the testimony.)

So transformative change involves incommensurable, epistemically inaccessible preference change from one self to another. We cannot distinguish acceptable transformative preference change from unacceptable transformative preference change merely by distinguishing adaptive reasoning or "sour grapes" preference formation from other types of preference formation.⁶⁶ Nor can we solve the problems of transformative choice simply by making a more careful comparison of the options for the agent. As a result, while Howard has made significant progress in developing our understanding of "I'll be glad I did it" reasoning, her solution to that problem fails to engage with the deeper problem of transformative choice that cochlear implant cases and other sorts of life-changing decisions raise.⁶⁷

Ruth Chang (2015) develops the connection between transformative experience and contemporary debates in ethics about reasons, self-constitution, normative character, and objective values. She describes my view of transformative choice as "event-based" transformative choice and contrasts it to "choice-based" transformative choice, distinguishing between the different types of transformation involved and their implications for practical decision-making. The transformation in choice-based transformative choice involves changes in one's normative character (the change could be minor or major), whereas event-based transformative choice involves major changes in one's epistemic capacities and personal preferences.

In her paper, Chang (i) explores the relationships between event-based transformative choice and choice-based transformative choice. She argues that event-based transformative choice poses no threat to decision theory because (ii) experiences like having a baby are not epistemically transformative, (iii) objective value rather than subjective value is the value of interest for transformative decision-making, and (iv) the possibility of radical personal transformation can be solved by standard approaches to rational choice.

⁶⁵ Similarly, if you are choosing for another person, such as a child, if what you chose at $t1$ formed her preferences at $t2$, the fact that her preferences at $t2$ are such that she prefers to have those preferences not justify your actions at $t1$. This relates to the discussion of choice ex ante versus choice ex post in my section 5.

⁶⁶ The classic text for this is Elster 1983.

⁶⁷ Howard, in discussion, emphasizes that her central project is to explore how deference to the testimony of others could play a role in morally and practically justifying our decisions. This feature of her project dovetails nicely with the concerns I raise in Paul 2014 about informed consent, disability, and testimony.

In reply to (i) I give a model for choice-based transformative choice, highlighting its distinguishing feature, mental commitment as a basis for character formation. I show how choice-based transformation differs from event-based transformation, and how it can be embedded into the structure of event-based transformation. I point out that the radical epistemic and personal transformation of event-based transformation can undermine the rationality of making mental commitments involved in choice-based transformative choice. In reply to (ii), I'll discuss a problematic assumption and direct readers to relevant literature. I'll reply to (iii) by explaining that subjective value is part of what grounds objective value in the cases of interest.⁶⁸ In reply to (iv), I'll explain how radical personal transformation creates problems for rational choice.

6.4.1 *Transformative Choice*

Chang distinguishes between what she describes as “event-based” transformative choice and “choice-based” transformative choice. In order to maximize the possibility for productive discussion, I will engage with Chang on her own terms, and assume that my view of transformative choice is close enough to the view that she describes as “event-based transformative choice” to make meaningful comparisons.

In event-based transformative choice, you choose to perform act *A*. On the simplest version, the choice to perform *A* is rational if performing *A* maximizes your expected subjective value. We can frame this in causal terms, as Chang's description suggests: the choice to perform *A* can lead to a transformative outcome *O* that is causally downstream from the choice. For simplicity, I'll assume *A* determinately causes *O*.

It is important to be clear about the causal structure here. On this view, some of the causal outcomes of making a transformative choice to perform *A* are events that transform you, and some are events of your transformation *O* (assuming you are in fact transformed). Given that causation is transitive, this simply amounts to saying that the choice causes your transformation, either directly or by being among the events in the causal chain leading to your transformation. It is certainly the case that in this sense, if you choose to act, you choose to transform yourself. In my introductory remarks, above, I characterized such transformative choices as “life-making” choices. As I put it in *Transformative Experience*:

when facing [transformative] big life choices, the main thing we are choosing is whether to discover a new way of living: life as a parent, or life as a hearing person, or life as a neurosurgeon, and so forth; that is, we choose to become the kind of person—without knowing what that will be

⁶⁸ To forestall confusion: to say “*x* is grounded in *y*” does not entail that *x* is entirely grounded in *y*. The language is similar to causal language: saying “*c* causes *e*” does not entail that *c* is the *only* cause of *e*.

like—that these experiences will make us into. (Paul 2014, 123)

Chang raises a series of interesting concerns about the rationality of self-formation and the problems for practical decision-making, and contrasts event-based transformative choice with choice-based transformative choice. To understand the contrast, we must first understand how choice-based transformative choice is supposed to be different from event-based transformative choice.

Chang describes choice-based transformative choice as the kind of choice where “you change who you are by the very making of a choice, not by some experience or event downstream from your choice” (2015, 239). How is this different from saying that performing *A* causes *O*?

The idea is that the choice-based transformative choice is a choice made *before* the choice that is the performance of act *A* leading to outcome *O*. Chang describes it this way: “When we choose in a thick sense, that is, by committing to an alternative, we create reasons for ourselves to choose it—our commitment is that in virtue of which we have a reason to do something” (242).

When you commit to an alternative, you choose to prefer that alternative over the others. After you commit to an alternative by (mentally) preferring one act-outcome (*A–O*) sequence over other possible sequences, you then choose again. This time, you choose to act in a way that is most likely to bring about that outcome, that is, you choose to perform *A*. To understand Chang’s sentence quoted above, then, we need to see that only the first use of “choose” in her sentence involves the “thick sense” of choice. This sort of choice is a mental decision to prefer an alternative, such as deciding to pursue a course of action. The second use of “choose” in the quote involves the “thin sense” of choice, a choice that is merely a performance of an act, such as physically performing act *A*.

This gives us an interpretation of the structure for choice-based transformative choice as a structure that is temporally prior to the causal structure involved in event-based transformation. Choice-based transformative choice involves a mental commitment to one act-outcome sequence over another, where mental commitments create reasons. The idea is that this structure is embedded in the early part of the causal sequence that ultimately leads to the choice to perform act *A*, and *A* in turn leads to outcome *O*.

Let’s flesh it out a bit more with reference to our paradigmatic example, choosing to have a child.

Start with an important distinction, drawn from the distinction between thick and thin choice, between committing to an outcome in the sense of *mentally deciding to endorse* performing act *A* that leads to outcome *O*, and committing to an outcome in the sense of *actually performing* act *A* in order to bring about *O*. We need to be clear about which type of commitment we are talking about, so call the first type “mental commitment” and the

second type of commitment “performative commitment.” Let’s take the outcome *O* to be the effect that is the final product: the transformed person. In our example, *O* is the outcome of becoming a parent and forming a parent-child attachment.

In choice-based transformation, first, you mentally commit to an act-outcome sequence from *A* to *O*. For example, you mentally commit to having a child, that is, you mentally commit to performing the act of having a child with the outcome of becoming a parent and forming a parent-child attachment. This mental commitment, by hypothesis, creates a will-based reason (*R*): so you now have a will-based reason to bring about *O* via *A*. In our example, you now have a will-based reason to have a child and become a parent with a parent-child attachment.

This new will-based reason constitutes a new normative character for you. This is what Chang wants to highlight when she says “So by choosing, we can create new reasons for ourselves, thereby transforming ‘who we are.’” In Chang’s sense, we change who we are by changing our normative character. Now that you have a reason to have a child, you have a new normative character, one which reflects your desire to become a parent. This is the structure that is distinctive of choice-based transformative choice.

The rest of the scenario for our paradigmatic case of choosing to have a child just involves event-based transformative choice.⁶⁹ Once you have a will-based reason *R* for bringing about *O* via *A*, you then make a performative commitment to bringing about *O* by choosing to *A*. So you then choose to perform act *A*, and bring about *O*. Continuing the example, because you have a will-based reason to have a child, you choose to have a child so that you will become a parent and stand in a parent-child attachment relation.

Once the structure of such transformative choices is worked out, we can see that choice-based transformative choice and event-based transformative choice involve two different types of change we can undergo when making big life choices. As a result, they raise different kinds of issues to address when we understand the possibilities for how we might construct ourselves via our choices.

Choice-based transformative choice concerns the rationality of practical decision-making given the way we understand and justify the mental commitments that create our reasons and normative character.⁷⁰ In choice-based transformation, your mental commitment creates new reasons that constitute a new normative character. The change in character need not

⁶⁹ We can also continue to add preliminary “choice-based transformative choice” structure in various scenarios: perhaps you mentally commit to mentally committing to the act-outcome sequence from *A* to *O*. Then you create a will-based reason *R** that is a reason for your will-based reason *R*. One interesting question concerns the way to understand and rationally justify this sort of regressive structure for will-based reasons. This question raises problems for debates in practical ethics over the nature of self-constitution and reasons.

⁷⁰ “We can now see how in choice-based transformative choices your choice can be both what transforms you and that in virtue of which you are transformed. In deciding whether to have a child, by hypothesis, the given reasons are on a par. You have the normative power

be a radical change. Your choice might be a small one, and so your new normative character might be pretty similar to your old one. Still, in this sense, your choice (your mental commitment) makes you who you are.

Event-based transformative choice concerns the possibility of radical epistemic change, the lack of a subjective value function, and the problems for decision-making when radical epistemic change is accompanied by radical personal change involving inconsistent preferences. In event-based transformation, you transform yourself by performing an act that causes a radical change in your epistemic capacities and your core personal preferences. In this much more dramatic sense your choice (your performative commitment) forms who you are.

While the types of transformation and self-formation are different in each kind of choice, the possibility of epistemic and personal transformation does raise a problem for choice-based transformative choice. In particular, if hard choices can result in outcomes that are epistemically and personally transformative, this can make the alternative act-outcome sequences “noncomparable” (in Chang’s sense) for an agent who mentally commits to one *A–O* sequence over another. This brings out how the possibility of transformative experience and transformative decision-making can have implications for debates about morality, self-constitution, and rationality.⁷¹

Return to the choice to have a child. If being a parent is epistemically and personally transformative, then the agent cannot grasp the subjective value of that outcome before she becomes a parent. This means that, at the mental commitment stage, the agent cannot evaluate the subjective value of the outcome of having a child. As a result, the value of the outcome of having a child is “noncomparable,” that is, the agent cannot compare it to the subjective value of the outcome of remaining childless. This case undermines Chang’s general thesis about how to understand mental commitment and choice-based transformative choice as a form of rational decision-making, since values for alternatives must be accessible and evaluable in order to guide reason and choice.

In other parts of her paper, Chang argues that epistemic and personal transformation create no problems for rational decision-making. I reject

to commit to one of the options or one of its features. You might commit to forming a parent-child attachment. That commitment just is choosing to have a child in the thick sense. That commitment then creates new will-based reasons for you to have a child, that is, your commitment is that in virtue of which you now have a new will-based reason to have a child. Your new will-based reason then interacts with your other, given, reasons and guides your choice in the thin sense. You may now have most all things considered reasons to choose to have a child. Your new will-based reason transforms you because it is a reason that determines your normative character. You are now the sort of person who has most all things considered reasons to have a child. Before the choice you were the sort of person for whom the reasons for having a child and remaining child were on a par. By choosing, you change the reasons that determine your normative character” (Chang 2015, 275).

⁷¹ In addition to the connection to Chang’s work, there are connections to Korsgaard 2009 and Kierkegaard 2006. Also see Barnes 2015b and Kemp 2015.

these arguments. Below, I'll discuss some interesting issues that come up in the discussion.

6.4.2 *Are Experiences Like Having a Baby Epistemically Transformative?*

Chang argues that the subjective values of the outcomes of transformative choices can be known, at least well enough, before the transformative experience occurs.

Maybe the experience of having a child falls under types that include the experience of being in a family, passing a kidney stone, having a pet, and so on. Since you've had experiences that fall under the same types before, you will know something about what it's like to have a child. (249)

The suggestion that passing a kidney stone (while your sister is visiting? as your cat looks on?) teaches you what it's like to give birth is, quite frankly, bizarre.⁷² While Chang's other objections to the possibility of epistemic transformation are many and far-ranging, they also miss the mark, largely because the views she attacks bear only a passing resemblance to my own views. As space is limited, readers interested in critical discussion of epistemic transformation should consult [Barnes 2015b](#), [Campbell 2015](#) and [Paul 2015a](#), as well as [Collins 2015](#), [Dougherty et al. 2015](#), [Kauppinen 2015](#) and my replies to these papers. [Sharadin \(2015, section 2\)](#), does an excellent job of characterizing the idea that it is the distinctive nature of the lived experience of having a child that is relevant to my argument concerning the rationality of choosing to have a child based on what it will be like to become a parent, not merely the experience of changing diapers, feeling tired, etc.

6.4.3 *Is Subjective Value the Type of Value of Interest?*

Chang argues that subjective value is not a suitable ground for the values of the relevant decision outcomes. Now, I've been at pains to point out above in my reply to Kauppinen that subjective value is not mere subjective feel. As I put it in [Paul 2014](#), it's the value of *lived experience*, and as such, should be able to ground, at least partly, the values involved in many big life decisions. The subjective value of a lived experience is not merely a matter of the phenomenal character of the internal characteristics of one's

⁷² Chang seems to radically underestimate the real-life epistemic difficulties here. Preliminary empirical results in psychology from the work of Josiah Nunziato and Fiery Cushman indicate that many people report having transformative experiences when they experience significant life events such as becoming a parent. In particular, they report that the experience changed them in ways that they could not foresee. In addition, further results suggest that people are over-confident about their ability to anticipate the changes that will occur in their beliefs, preferences, desires, and values as a result of a transformative experience. (I thank Fiery Cushman for discussion.) For an interesting and relevant example of the kinds of complications and massive difficulties involved in gaining and regaining sight see [Sacks 1993](#).

inner life. It's a richer value, a value that includes what it's like to live in a particular set of circumstances, to live one's life in a particular way, or to "live an outcome." (For further discussion of how subjective value extends past one's internal mental life, see [Campbell 2015](#) and [Paul 2015a](#).)

Chang thinks that what really matters for transformative choices are what she describes as "objective values" and "objective goods," not subjective values. She uses the example of Mike May, who was blind but regained his sight through an operation:

What matters in May's choice about whether to see again are not only the objective and subjective values of the *experience* of seeing but also the objective *goods* (which we can characterize in terms of events) he will have in his life if he is sighted. Indeed, it makes sense to think that it was not the experience of seeing that primarily transformed him but other events, like communing with his wife over a beautiful sunset, responding to visual feedback from his children, and learning new skills that gave him greater opportunities that did the transforming work. ([Chang 2015](#), 262)

As an objection to the need for assessments of subjective value this is puzzling. For of course goods like "communings with his wife over a beautiful sunset" are the ones that matter. But a good like this depends at least partly on the nature of May's lived experience, and the subjective values of the experiences Chang describes are precisely what May cannot evaluate before he has his operation.⁷³ As I'd put it, for May, a grasp on the subjective value of regaining his sight is necessary to assess whether an outcome like communing with his wife over a beautiful sunset is objectively good—and if it is objectively good, just how good it is.

Some of Chang's arguments about our grasp of the relevant values concern testimony, and her objections reflect those that are raised by other contributors to this volume. I discuss the relationship between experience and subjective value in my reply to [Kauppinen 2015](#), and discuss the limitations of testimony in determining subjective value in my discussion with [Dougherty et al. \(2015\)](#), above, with [Harman \(2015\)](#), [Howard \(2015\)](#), and in my exchange with [Richard Pettigrew \(2015; 2015a\)](#).

6.4.4 *Does the Possibility of Radical Personal Transformation Challenge Decision Theory?*

Chang suggests that decision theory can manage the cases involving personally transformative experiences that [Edna Ullmann-Margalit \(2006\)](#) and I

⁷³ We are assuming, by hypothesis, that even though May once had sight, before having his operation to restore his sight, he lacked the capacity to represent events involving seeing things like sunsets.

describe. She claims that the problem can be solved simply by predicting and assessing future preferences or by employing “a master utility function” that could order consecutive sets of preferences with respect to any given choice. But, of course, this misdescribes the situation. Our examples of personal transformation crucially concern *discontinuous* preferences, including discontinuous higher-order preferences. These are cases where no master utility function or prospective resolution is rationally available.⁷⁴

It is worth setting the context with some familiar examples. Jon Elster considers a case where Ulysses knows his future self will be temporarily irrational. In such a case, Elster argues, it is rational for Ulysses to bind his future (irrational) self (Elster 1979). Derek Parfit’s young Russian nobleman values wealth redistribution, but knows that when he is old, he will value keeping his money instead. By assumption, the preferences of the young Russian nobleman are discontinuous with the preferences of the old Russian nobleman. Parfit suggests that the solution is for the young Russian nobleman to bind his future self, just as Ulysses did, even though the nobleman’s future self is not irrational.⁷⁵

If we have rational grounds for privileging the preferences of the current self over the preferences of the future self, perhaps because the current self and the future self share higher-order preferences (or there is some independent reason that the future self’s preferences are rationally disqualified), then the current self can rationally choose to bind the future self consistent with her current higher-order preferences.

However, this solution is untenable in transformative contexts such as choosing to have a child: it cannot provide a rational guide for life’s transformative decisions.⁷⁶ In transformative decisions like choosing to have a child, certain core preferences of the current self are radically discontinuous with preferences of the future self, and there is no independent basis for disqualifying these preferences of the future self. Moreover, because of the epistemic inaccessibility of the preferences of the future self, the current self cannot (imaginatively) prospectively assess the lived experience of the future self to form a higher-order preference that would be consistent across the selves at different times.⁷⁷

There is another implication in the neighborhood. If morality is bound up with rationality, and the decision for one’s future self concerns a morally transformative experience, then the problem for rational decision-making, especially for rational decisions concerning the construction of one’s future

⁷⁴ Moreover, in my cases, we lack epistemic access to our future preferences. For discussion of the issue, see Briggs 2015 and Pettigrew 2015 and my replies.

⁷⁵ Korsgaard (2009) challenges this.

⁷⁶ Unless, as I argue, the current self chooses solely on the basis of preferring to discover new preferences, that is, on the basis of the preference to replace the current self with a new self. (Or, presumably, to end the existence of the current self, independently of the question of replacement.)

⁷⁷ See Briggs 2015 for a different response.

self, becomes a problem for moral decision-making and the construction of one's future moral perspective.

Korsgaard (2009) argues, in circumstances where core preferences will be radically transformed, your current self must either take your future self's preferences into account, or regard your future self as irrational. In my cases of transformative choice, you cannot take your future self's preferences into account. First, you lack first-personal epistemic access to your future self's preferences. And second, you cannot simply prefer to prefer your future self's preferences, because they conflict with your current preferences, including your higher-order preferences. Third, being told by an expert what your future self would prefer you to do now is not sufficient for you to prefer that future self's preferences, because without first-personal imaginative understanding of this future self, you are entitled to regard her as irrational (or alien).

Thus, the problem remains: how can it be rational to choose to have a child? Or, as one might put it, how can it be rational for you to make yourself into someone you regard as irrational?⁷⁸

6.5 Empirical Research and Choosing to Have a Child

There is a certain sort of easy reply to the problem of transformative decision-making. It is to say that a person should simply replace her introspection with scientific evidence that will tell her what to expect.⁷⁹ Nathaniel Sharadin (2015) argues for a version of this reply, arguing that you can use currently available empirical research to predict what it would be like for you to become a parent, and this will solve the problems with the transformative decision to become a parent. His reply fails. Why?

First, there are practical difficulties. As I discuss in [section 4](#), above, contemporary psychological work is simply not yet advanced enough for us to use it to make sufficiently accurate predictions in our own case about how we'd respond to having a child. Once introspection is set aside, without suitably precise empirical information about my own particular, individual response to having a child, I'm left adrift. All I have is highly general, incomplete, empirical information that I'm somehow supposed to interpret in a way that applies to my own particular case. In a high-stakes case like this, where I am making one of the most important, irreversible, momentous, and personal decisions of my life, currently available psychological data cannot provide a satisfying replacement for introspection.

Second, even if we do have sufficiently complete, detailed empirical results, there are deeply philosophical interpretive difficulties with the testimony grounding the evidence. Almost without exception, one of the most serious problems is that the empirical results measure the preferences and

⁷⁸ In [Paul 2014](#), I propose a solution involving revelation.

⁷⁹ We saw a version of this reply in [Harman 2015](#), where the suggestion is that we should simply rely on "reliable testimony."

the satisfaction of *already transformed* individuals.⁸⁰ As I have discussed in detail in [section 5](#) above, the epistemic perspective and preferences of the agent after she is transformed may be discontinuous with her epistemic perspective and preferences when she is making the decision, raising questions about how she is to understand and interpret the evidence.

Sharadin (2015) contests both claims, but his argument focuses primarily on the first difficulty: the use of current psychological and sociological data to make the transformative decision to become a parent. He grants that introspection about the nature of the experience can fail, but thinks a person can simply replace introspection with current empirical research to discover what it would be like for her to have a child.⁸¹

[P]rospective parents cannot rationally decide to have a child by reflecting on the phenomenal character of that experience: it's in principle epistemically inaccessible to them. But this does not mean that prospective parents cannot rationally decide to have a child by reflecting on what it is like to have a child. It just means they have to take a somewhat circuitous route: prospective parents must reflect on the non-phenomenal features of the experience, on what they themselves are like, and on the principles that link how they are to how the experience is likely to affect them. (450)

According to Sharadin, contemporary social and psychological science gives us these linking principles, and a person can use them to “reasonably expect” the valence of what it will be like for her to become a parent. From this, he concludes that a person can indeed rationally choose to become a parent based on what it will be like for her.

There are two serious problems with his argument.

The first problem is that Sharadin claims that I deny the existence of such linking principles. But I do not deny the existence of such linking principles. I am happy to grant that there might exist such psychological “laws” for individuals. It's discovering them that's the problem.

So let's assume that linking principles exist. The second, much more important problem with Sharadin's argument is that he thinks it is manifestly obvious that we know what many of these linking principles are, and that an individual can and should use them to determine what the valence of her outcomes will be.

⁸⁰ There are also problems involving average effects, the fundamental identification problem, and reference class worries. These problems are significantly more serious when introspection is unavailable. See [Paul 2014](#) and [2015a](#).

⁸¹ I found Sharadin's discussion in section 2 of his paper, of the distinctive nature of what it's like to have a child, to be thoughtful and interesting, and his discussion of the role of swamping and the relevance of distinctive subjective values in the argument for epistemic transformation is right on target. Section 3 is where the problems begin.

[T]here manifestly *are* such linking principles, and we know what some of them are. For just one example, depression on the part of either parent, but especially maternal depression, is linked to both affective and behavioral disorders on the part of children (Lovejoy et al. 2000; Tan and Ray 2005). And parents of affectively or behaviorally disordered children report significantly higher rates of stress and lower levels of subjective well-being—as good a measure as any of the valence of the phenomenal character of their experience of what it is like to have a child (Tan and Ray 2005, 77). (Sharadin 2015, 449)

He goes on to argue that

This should come as no surprise at all: what people are like helps determine how things turn out for them. And, thanks to years of psycho- and sociological research, we can often safely predict how things will turn out for an agent given enough psycho- or sociological information about them. Of course, the situation is no different when it comes to ourselves than it is in the case of others. Or at least, it is not relevantly different. Just as I can know that, given that some agent is depressed, the phenomenal character of her experience of having a child is unlikely to be positive, I can know of myself that, given I am depressed, the phenomenal character is unlikely to be positive. And so, *ceteris paribus*, I can safely predict that it would be unwise, just now at least, for me to have that experience.⁸² (449)

Unfortunately, however, Sharadin has misunderstood the implications of the work he is citing, for Tan and Ray (2005) and Lovejoy et al. (2000) don't give us anything like what an agent would need to in order to reasonably predict, in her own individual case, what it would be like for her to have a child. (By extension, they do not give us the linking principles he thinks they provide.) This is no surprise: as I noted in Paul 2015b, the relevant psychological and sociological science is not yet complete enough for an individual to discover her own personal linking principle or even a reasonable approximation thereof.

A closer look at the papers will be instructive: it will demonstrate the perils of relying on naïve interpretations of empirical work in an attempt to shrug off the implications of transformative experience for decision-making. In particular, a closer look will bring out first, how easy it is for nonexperts to misunderstand what empirical work actually tells us, and second, just how difficult it can be to take empirical results and apply them to your own case.

⁸² He acknowledges “there might be countervailing reasons . . . to expect it will be positive” but “The point is just that such expectations are sometimes warranted” (449).

The Tan and Ray paper is a small study done in Malaysia. The authors have a matched sample of depressed and non-depressed children, and ask whether parents of depressed children have a harder time. The answer seems to be yes: “Parents of depressed children reported higher parenting stress and were more likely to perceive their children as ‘difficult’” (Tan and Ray 2005, 76). But note: this study does not give you information about your chances of having a child that will be depressed. Moreover,

[u]nivariate analysis demonstrated a relationship between children’s depression and maternal depression but not with paternal depression. This became non-significant with multivariate analysis, implying that maternal depressive symptoms may have been due to caring for a depressed, thus ‘difficult’ child. (76)

In other words, a prospective parent could worry that depressed mothers have depressed children. But while the data show a correlation between the two, simple further analysis with a few controls made the association disappear. This suggests that it might be the child that makes the mother depressed, not the other way around.

So a closer look at the Tan and Ray paper provides no support for the claim that it provides knowledge of linking principles a person could use to predict her outcomes of what it would be like, if she were depressed, for her to become a parent.

The Lovejoy et al. 2000 paper is a much more substantial review. It is a widely cited meta-analysis seeking “to assess the strength of the association between depression and parenting behavior” (561). It’s a typical and responsible example of its type, covering more than twenty years of peer reviewed studies of different kinds on the relationship between maternal depression and parenting behavior.

Again, note that this review does not speak at all to the question of who is likely to become depressed should they have a child. Lovejoy et al only look at studies of people who are *already* mothers and, moreover, at mothers who are *already* depressed.

With this in mind, consider the findings in Lovejoy et al. 2000. For the sake of argument, let us simply assume that, somehow, as a prospective parent, you already know that you will be depressed after you have your baby. This would be extremely difficult to establish in practice in most cases, and making this assumption rather misses the point of the entire discussion of transformative experience. But as the empirical evidence Sharadin cites would be otherwise completely irrelevant, it seems charitable to grant the assumption. What then can we “safely predict” from the data presented?

Lovejoy et al’s review asks, given that you are a depressed mother, how are you likely to behave toward your child? Do depressive mothers act more negatively towards their children, are they dissociated from them, or are they positive towards them? Lovejoy et al suggest there is indeed an

observed relationship between maternal depression and parenting behavior. But it is not a simple one.

The association between depression and parenting was manifest most strongly for negative maternal behavior and was evident to a somewhat lesser degree in disengagement from the child. The association between depression and positive maternal behavior was relatively weak, albeit significant. (561)

This result is summarized in Table 3 of the paper (Lovejoy et al. 2000, 579). They find a positive and statistically significant association between maternal depression and negative behavior. They also find a positive and significant association between maternal depression and disengaged behavior. And finally they find a positive and significant association between maternal depression and positive behavior! (The effect sizes vary. The association with negative behavior is the largest. The association with positive behavior is about two and a half times smaller, but still positive. It is also significant according to conventional standards.)

If it seems strange to you that all three behaviors could be positively associated with maternal depression, bear in mind this is a meta-analysis of many studies. Positive and negative behaviors are not measured on a single scale with positive on one end, negative on the other, and dissociated in the middle. Instead they are measured separately through the observation of many behaviors.

From a first-personal point of view, this means that—given that we have granted that you know with certainty that you will be depressed after you have your baby—each of the measured behavioral outcomes remains a live possibility for you. It is a possibility just in the important but weaker sense that you might be an “outlier” with respect to a single general tendency in the population, but also in the specific sense that the review finds all three patterns are in fact observed as significant general tendencies. This means that acting on the basis of the best evidence you have, you, as a depressed person, are left largely unsure about what sort of behavior you would be likely to show to your child, if you had one. While you can certainly attach credences to different possible outcomes based solely on this very general information, it is hardly specific enough to count as giving you evidence (much less knowledge) about your own possible outcomes, let alone a workable “linking principle” for a decision at the individual level.

So how, then, is our prospective decision-maker to “safely predict” how she is likely to behave? At the outset, the authors issue a general caution:

It is important to note, however, that the association between depression and child adjustment problems may not be causal. Child behavior problems could, for example, contribute to the development of maternal depression. It is also possible that a third variable is causally related to both

maternal depression and child adjustment problems. . . . [T]here are a number of associated features of depression that may account for the relationship between maternal depression and childhood difficulties. These include substance abuse, personality disorders, and marital discord. (562)

This is an important caveat, and the fact that it is a routine one in summaries of observational studies should not tempt us to simply set it aside. As we move on to the details, we see, as is typical of reviews of this sort, that the authors make a careful and cautious interpretation of the observed patterns:

Our statistical analyses of 46 observational studies demonstrated a moderate association between maternal depression and parenting behavior in the domain of negative behavior, a small to moderate effect in the domain of disengaged behavior, and a small effect for positive interactions. Thus, depression appeared to be associated most strongly with irritability and hostility toward the child, to be associated to a somewhat lesser degree with disengagement from the child, and to have a relatively weak association with rates of play and other active and pleasant social interactions. However, there was marked variability in the effect sizes obtained in each domain of parenting behavior, only some of which could be explained by the moderators we included in our analyses. (583)

So although some general patterns are evident, there is more than one significant tendency. What we should conclude from this is that, even with twenty years of data, things are a long way from being fully explained. Moreover, the available covariates are both relatively few in number and relatively limited in their explanatory power. The authors also express some (again, responsible and quite typical) concern about how different behaviors were coded and assessed across studies. They note for example that

negative coercive behaviors and positive behaviors were coded fairly evenly across the studies, regardless of child age; however, disengaged behavior was assessed primarily in mothers of young children. . . . Because the behaviors coded in the original studies were more varied for the mothers of very young children, it lends the appearance that the vestigial parenting problems associated with depression are more pervasive for infants and toddlers. We believe this is more than an artifact of the coding system and reflects the dependency of infants and very young children on their

caretakers to initiate interaction, and maintain contact that is coordinated with the child's affect and behavior. (584)

Heterogeneity in the measurement instrument, and its possible correlation with outcomes of interest (e.g., the age of the child), make interpretation even more difficult. So, for example, the authors note that:

Although certainly not conclusive, the obtained pattern of age differences argues against strong child effects in the development of depression and parenting difficulties.

And:

Although our findings do not resolve questions of causality in the relationship between depression and parenting, they also do not suggest that the majority of parenting difficulties of depressed mothers originate from individual differences in child behavior. (585)

Because the authors are specifically interested in the effects of depression—considered not simply as a folk category but as a well-defined clinical condition—they also note a potential theoretical issue arising from the findings:

With respect to the pattern of effect sizes, our analyses demonstrated that the largest effects occurred for negative/coercive behavior; however, involvement, sensitivity, and pleasant social interactions would logically seem to be most sensitive to depressive symptoms. Relying solely on the diagnostic criteria for depression, one would predict that depressed mothers would have the most difficulty in the domain of engagement. . . . Thus, the defining characteristics of depression would lead us to expect the effect sizes to be largest for disengaged and positive behavior and smallest for negative/coercive behavior. In contrast, we found the effect size to be largest for negative behavior and smallest for positive behavior. This finding, which is not consistent with predictions based on the symptoms of depression, suggests that some of the parenting problems observed among depressed mothers may be associated with negative affectivity, and that these same parenting difficulties may occur among women with other emotional problems and general psychological distress. (587)

Here we see a pattern of reasoning—again, very common in research of this kind—where the results prompt a reevaluation of the original theory rather than simply establishing a “finding” with a clear recommendation attached.

The authors suggest that instead of thinking in terms of depression,

conceptualizing depressive parenting problems in terms of disturbances in positive and negative affect leads to predictions somewhat more consistent with the findings of our meta-analysis and provides a useful theoretical model from which to interpret our results. (587)

Note that the concern *is not to develop or pinpoint advice to particular sorts of individuals*, let alone for individuals in any fine-grained sense. Instead the concern is to try to square the observed pattern of results with a general claim about the nature of an association *across a population*.

The paper closes with a call for further research, and more caution:

In summary, our analyses suggest that the strength of the association between depression and parenting behavior varies as a function of the type of behavior observed. Effects were strongest for negative/coercive behavior and least strong for positive interactions. However, there was significant variability in effect sizes within each category of behavior and many of the effect sizes for individual studies did not differ from zero. Although our findings suggest that age of child, socioeconomic status, and timing of depression account for some of the differences between studies, significant heterogeneity in effect sizes remained. . . . Explication of the variability of effects remains a critical task for understanding the developmental risks associated with maternal depression. . . . The results of our analyses are consistent with the hypothesis that parenting behaviors are a component of the risk associated with living with a depressed mother. . . . However, further research is needed to more fully understand which children of depressed mothers are most likely to be exposed to inadequate parenting. (588)

This should make it abundantly clear that the research in the papers cited by Sharadin has nothing to say to the prospective parent worrying about whether having a child will make them depressed. This is because they only study the behavioral consequences, positive and negative, for people who are already parents and already depressed. And, as should also be clear at this point, even a depressed individual will not find any “linking principles” that will support inferences about what it will be like for *her* to become a parent.

However, in a deeper sense, the reviews do cast light on the role of research like this in making potentially transformative decisions, and this is why I have discussed them at such length. We find at least three kinds of illumination. First, the careful analysis and discussion on the part of the authors of these studies shows just how difficult it can be to make a clear causal analysis in these cases, and how reluctant responsible researchers are

to go beyond the data to make simple recommendations to specific people about how they should choose based on the evidence.

To be sure, there are some cases where things are easier. If you and your partner are tested for Huntington's Disease, then you will learn some very specific and relevant facts about your own future and the chance that any future child of yours will have the disease. But many or most relevant psychological aspects of your possible future as a parent are not as causally specific or empirically testable as this.

Second, even very good, responsible sociological and psychological research—of the sort carried out and reported in Lovejoy et al—can take an extremely long time to reach its conclusions, relative to the decision windows of prospective parents. The questions raised in the Lovejoy et al review could well prompt a further twenty or thirty years of study that might yield a round of real but slow progress. Science advances. But for many who have to choose now, or soon, it does not advance nearly fast enough. And so even if we, collectively, eventually arrive at a well-established and relevant body of knowledge, being able to “safely predict” outcomes of interest is, for a huge range of cases, a very long way off. There is no point in grandly referring to “years of psycho- and sociological research” that either does not give you the information you need now, or that might provide it to you in a more appropriately fine-grained fashion a few decades after your death.

Third, as I said at the beginning, although it is a model of responsible meta-analysis, Lovejoy et al's article is in fact strictly irrelevant to the person facing a potentially transformative choice. Sharadin might object that the cited papers were merely “just one example” meant to serve as a sort of placeholder for a vast body of well-established, properly-validated, causally impeccable, internally consistent social-scientific findings of direct relevance to individuals facing the particular choices they are interested in. I do not think this body of knowledge exists in this form.

In an interesting way, Sharadin's discussion mirrors the epistemic attitude of most people who face possibly transformative decisions in their own lives. Such people often have strong intuitions about what they want; they have a reasonable belief that research findings should have a role to play in helping them choose; and they have a conviction that there must be a rational course of action that combines their feelings with “the data.”

But where the research, especially in these areas, is careful, complex, controversial, slow-moving, and concerned mostly with tendencies at the level of whole groups, ordinary people need to make individual-level decisions, and they need to make them *now*. And so they sit at their computer and cast their line into a sea of scientific research. They reel in a few studies and read the abstracts. They fail to grasp what the research can actually establish, they glide over the caveats inserted by the authors, and they convince themselves that with these papers in hand they can—*ceteris paribus!*—“safely predict” their own futures. The result is a parody of

rational choice, a pretense of rationally justified decision-making rather than a clear-eyed step into an acknowledged unknown.

L. A. Paul

E-mail : lapaul@unc.edu

References:

- Barnes, Elizabeth. 2014. "Fundamental Indeterminacy." *Analytic Philosophy* 55 (4): 339–362. <http://dx.doi.org/10.1111/phib.12049>.
- Barnes, Elizabeth. 2015a. "Social Identities and Transformative Experience." *Res Philosophica* 92 (2): 171–187. <http://dx.doi.org/10.11612/resphil.2015.92.2.3>.
- Barnes, Elizabeth. 2015b. "What You Can't Expect When You Don't Want to be Expecting." *Philosophy and Phenomenological Research* 91 (3).
- Barnes, Elizabeth. Forthcoming. *The Minority Body*. Oxford: Oxford University Press.
- Briggs, Rachel. 2015. "Transformative Experience and Interpersonal Utility Comparisons." *Res Philosophica* 92 (2): 189–216. <http://dx.doi.org/10.11612/resphil.2015.92.2.7>.
- Campbell, John. 2015. "L. A. Paul's *Transformative Experience*." *Philosophy and Phenomenological Research* 91 (3).
- Carr, Jennifer. 2015. "Epistemic Expansions." *Res Philosophica* 92 (2): 217–236. <http://dx.doi.org/10.11612/resphil.2015.92.2.4>.
- Cartwright, Nancy. 2011. "The Art of Medicine: A Philosopher's View of the Long Road from RCTs to Effectiveness." *The Lancet* 377: 1400–1401. [http://dx.doi.org/10.1016/S0140-6736\(11\)60563-1](http://dx.doi.org/10.1016/S0140-6736(11)60563-1).
- Chang, Ruth. 2015. "Transformative Choices." *Res Philosophica* 92 (2): 237–282. <http://dx.doi.org/10.11612/resphil.2015.92.2.14>.
- Collins, John. 2015. "Neophobia." *Res Philosophica* 92 (2): 283–300. <http://dx.doi.org/10.11612/resphil.2015.92.2.6>.
- Dougherty, Tom, Sophie Horwitz, and Paulina Sliwa. 2015. "Expecting the Unexpected." *Res Philosophica* 92 (2): 301–321. <http://dx.doi.org/10.11612/resphil.2015.92.2.5>.
- Elster, Jon. 1979. *Ulysses and the Sirens*. Cambridge: Cambridge University Press.
- Elster, Jon. 1983. *Sour Grapes: Studies in the Subversion of Rationality*. Cambridge: Cambridge University Press.
- Harman, Elizabeth. 2009. "I'll Be Glad I Did It' Reasoning and the Significance of Future Desires." *Philosophical Perspectives* 23: 177–199. <http://dx.doi.org/10.1111/j.1520-8583.2009.00166.x>.
- Harman, Elizabeth. 2015. "Transformative Experience and Reliance on Moral Testimony." *Res Philosophica* 92 (2): 323–339. <http://dx.doi.org/10.11612/resphil.2015.92.2.8>.
- Howard, Dana Sarah. 2015. "Transforming Others: On the Limits of 'You'll Be Glad I Did It' Reasoning." *Res Philosophica* 92 (2): 341–370. <http://dx.doi.org/10.11612/resphil.2015.92.2.9>.
- Johnston, Mark. 2006. *Perceptual Experience*. Edited by T. S. Gendler and John Hawthorne. Oxford: Oxford University Press.
- Kauppinen, Antti. 2015. "What's So Great about Experience?" *Res Philosophica* 92 (2): 371–388. <http://dx.doi.org/10.11612/resphil.2015.92.2.10>.

Acknowledgements I'd like to thank Nomy Arpaly, Elizabeth Barnes, Paul Bloom, Rachael Briggs, Jennifer Carr, John Collins, Fiery Cushman, Elizabeth Harman, Kieran Healy, Dana Howard, Jonathan Jacobs, Antti Kauppinen, Matt Kotzen, Enoch Lambert, Rachel McKinnon, Sarah Moss, Josiah Nunziato, Cass Sunstein, Julia Staffel, and Greg Wheeler for discussion. This research is part of the Experience Project, and was supported with a grant from the John Templeton Foundation. The opinions expressed in this publication are those of the author's and do not necessarily reflect the views of the John Templeton Foundation.

- Kemp, Ryan. 2015. "The Self-Transformation Puzzle: On the Possibility of Radical Self-Transformation." *Res Philosophica* 92 (2): 389–417. <http://dx.doi.org/10.11612/resphil.2015.92.2.11>.
- Kierkegaard, Søren. 2006. *Fear and Trembling*. Edited by C. Stephen Evans and Sylvia Walsh. Cambridge: Cambridge University Press.
- Korsgaard, Christine. 2009. *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Krishnamurthy. 2015. "We Can Make Rational Decisions to Have a Child: On the Grounds for Rejecting L. A. Paul's Arguments." In *Permissible Progeny*, edited by Richard Vernon, Sarah Hannan, and Samantha Brennan. Oxford University Press.
- Lewis, David. 1989. "Dispositional Theories of Value." Supplementary Volume. *Proceedings of the Aristotelian Society* 63: 113–137.
- Lovejoy, M. Christine, Patricia A. Graczyk, Elizabeth O'Hare, and George Neuman. 2000. "Maternal Depression and Parenting Behavior: A Meta-analytic Review." *Clinical Psychology Review* 20 (5): 561–592. [http://dx.doi.org/10.1016/S0272-7358\(98\)00100-7](http://dx.doi.org/10.1016/S0272-7358(98)00100-7).
- McKinnon, Rachel. 2015. "Trans*formative Experiences." *Res Philosophica* 92 (2): 419–440. <http://dx.doi.org/10.11612/resphil.2015.92.2.12>.
- Moss, Sarah. Unpublished. "Probabilistic Knowledge."
- Paul, L. A. 2014. *Transformative Experience*. Oxford: Oxford University Press.
- Paul, L. A. 2015a. "Transformative Experience: Replies to Pettigrew, Barnes, and Campbell." *Philosophy and Phenomenological Research* 91 (3).
- Paul, L. A. 2015b. "What You Can't Expect When You're Expecting." *Res Philosophica* 92 (2): 149–170. <http://dx.doi.org/10.11612/resphil.2015.92.2.1>.
- Paul, L. A. Unpublished. "Preference Capture."
- Pettigrew, Richard. 2015. "Transformative Experience and Decision Theory." *Philosophy and Phenomenological Research* 91 (3).
- Sacks, Oliver. 1993. "To See and Not See." *The New Yorker*.
- Sharadin, Nathaniel. 2015. "How You Can Reasonably Form Expectations When You're Expecting." *Res Philosophica* 92 (2): 441–452. <http://dx.doi.org/10.11612/resphil.2015.92.2.2>.
- Stanley, Jason. 2015. *How Propaganda Works*. Princeton, NJ: Princeton University Press.
- Sunstein, Cass. 2014. "The Limits of Quantification." *California Law Review* 106 (2): 1369–1422.
- Tan, Susan and Joseph Ray. 2005. "Depression in the Young, Parental Depression and Parenting Stress." *Australasian Psychiatry* 13 (1): 76–79. <http://dx.doi.org/10.1080/j.1440-1665.2004.02155.x>.
- Ullmann-Margalit, Edna. 2006. "Big Decisions: Opting, Converting, Drifting." *Royal Institute of Philosophy Supplement* 58: 157–172. <http://dx.doi.org/10.1017/S1358246106058085>.
- Van Fraassen, Bas. 1999. "How is Scientific Revolution/Conversion Possible?" *Proceedings of the American Catholic Philosophical Association* 73: 63–80. <http://dx.doi.org/10.5840/acpapro19997313>.
- Velji, Muhammad. 2015. "Change Your Look, Change Your Luck: Religious Self-Transformation and Brute Luck Egalitarianism." *Res Philosophica* 92 (2): 453–471. <http://dx.doi.org/10.11612/resphil.2015.92.2.13>.
- Vermeule, Adrian. 2013. "Rationally Arbitrary Decisions (in Administrative Law)." *Harvard Public Law Working Paper No. 13–24*. <http://dx.doi.org/10.2139/ssrn.2239155>.
- Zelcer, Mark. 2015. "Conscientious Objection and the Transformative Nature of War." *Journal of Military Ethics* 14 (2): 118–122. <http://dx.doi.org/10.1080/15027570.2015.1070035>.