ADAPTIVE PREFERENCES AS AUTONOMY DEFICITS

Rosa Terlazzo

Introduction

What are adaptive preferences? As with so much in philosophy, it depends on who you ask. But most philosophers writing about adaptive preferences can at least agree on this: adaptive preferences are preferences that you hold more because your option set is bad or limited than because of the good-making features of the objects you prefer.\(^1\) In other words, you want what you do because it was the best option available to you, and you would have wanted something else instead if you’d had a better or larger set of options.

Beyond that, theories of adaptive preferences don’t agree on much, and many of those who discuss adaptive preferences don’t really have a theory of adaptive preferences at all – instead, they offer examples of preferences that are intuitively objectionably adaptive to motivate an account of what should be done with respect to them. In this chapter, I’m not going to try to answer the extremely fractious question of what precisely adaptive preferences are; instead, I’m going to investigate why one might want to think about them in terms of autonomy, and lay out a framework for thinking about which subset of non-autonomous preferences the adaptive ones might be.\(^2\)

I’ll begin by developing an understanding of adaptive preferences via examples of the kinds of cases that all accounts of adaptive preferences will agree on. Then I’ll discuss what a preference’s adaptiveness might license, and why (lack of) autonomy seems to offer a compelling explanation of what makes adaptive preferences objectionable. From there, I’ll turn to clarifying which subset of non-autonomous preferences adaptive preferences might be, focusing first on the non-autonomous preferences that everyone would agree are not adaptive preferences, and then discussing the kinds of autonomy-relevant features that might reasonably divide adaptive preference theorists.

1. Adaptive Preferences and Autonomy

Let’s begin with some intuitively compelling examples of adaptive preferences:

Happy slave: The Happy Slave was born into slavery and is glad she is enslaved. She is relieved that she doesn’t have to make any major choices about her life, and that she will
Adaptive Preferences as Autonomy Deficits

always have food to eat and a roof over her head. She sees the punishments that come with slavery as a good incentive to be an industrious and committed worker. When her fellow enslaved people protest their conditions, she wonders, “What’s so great about freedom anyway?”

Resigned laborer: The Resigned Laborer works long and punishing hours in a brick kiln for very little money – and indeed, for even less money than her male co-workers. The work is monotonous and hard and her body is ravaged by it, but when approached by a labor organizer she says “It’s not so bad; life is hard, and men are always paid more. That’s just the way it is. No point in complaining about it.”

Deferential wife: The Deferential Wife defers to her husband in all major decisions. She believes that men should lead families, and that women should care for them and support them – and of course raise the children. When she wants something that her husband doesn’t want her to have or do, she tells herself “I don’t know what is good for me anyway” or “It’s the natural order of things for men to make these decisions – it would be proud and vain for me to think it’s my place to do it.”

Manly man: The Manly Man will get angry at the world, but he would not be caught dead crying over anything except the death of his dog or his team losing a major tournament. He both accepts and demands emotional support from his mother and his wife, but he doesn’t talk about his feelings, and he would certainly never see a therapist. When he feels sad or lonely, he gets angry or pushes his feelings down deep and thinks to himself “Stop being a pussy – real men aren’t weak like this, and they don’t cry about their feelings to others.”

Assuming that the person in each of these examples is sincere in their preferences, a few things unify them. Most obviously, all of them are in situations that seem intuitively to be quite bad for them. Neither the Deferential Wife nor the Happy Slave gets to make major choices about her life, and both the Resigned Laborer and the Happy Slave must do back-breaking labor whose fruits go to someone else. Both the Deferential Wife and the Resigned Laborer are in positions where men have significant and undeserved power or advantages over them simply by virtue of their gender, and both the Manly Man and the Deferential Wife are constrained by sexist norms that prevent them from having their basic needs and desires met.

Second, the people in all four of the examples endorse their situations to at least some degree. The Happy Slave doesn’t want the freedom that her fellows do, the Resigned Laborer doesn’t think the situation is worth protesting, and both the Deferential Wife and the Manly Man seem to have internalized and approve of the standards that constrain them. To be sure, we could fill in the details in each case so that the endorsement was more or less strong – perhaps the Deferential Wife very happily endorses gender hierarchy, for instance, or perhaps she just thinks it’s both as unpleasant and as unavoidable as unclogging stopped up toilets. But in each example, our subject falls somewhere between enthusiastic defense of their situation, and thinking that nothing about their situation need change.

And finally, in each example, the subject’s preferences seem to be the product of their situation. The Resigned Laborer would almost certainly take a higher-paying or less arduous job if it were actually offered to her, and it’s hard to imagine that the Happy Slave would have chosen slavery if she had not been born into it. Both the Manly Man and the Deferential Wife seem to be products of pervasive and rigidly endorsed gender norms. In each case, it seems highly likely that if the subject would have had a larger or better set of options, they would have a very different – and intuitively better – set of preferences about their lives.
Accounts of adaptive preferences, then, seem to be unified by the idea that finding yourself in bad or limited circumstances can cause you to endorse things that are bad for you, and that you would reject if those circumstances had not arisen.

To understand why it might be helpful to think about adaptive preferences as autonomy deficits, it’s helpful to think about why we have the concept of adaptive preferences in the first place. Generally, the concept is used to undercut the kind of first-personal authority we usually grant people over the extent to which their lives are good for them. Absent reasons to think otherwise, we generally defer to people when they tell us what makes them happy and how it would be good for their lives to go. You don’t like spicy food? You want to be a dentist? You love opera? You prefer leisure time to wealth? In each of these cases, most of us would tend to accept that these kinds of lives are good for the people these statements describe, even if we don’t think we ourselves would like that life much at all. The concept of adaptive preferences gives us one way of justifying our judgments in cases where we are not willing to grant that other people’s judgements about their own lives are sound. They help to explain why someone would want something that wasn’t good for her – namely, because she found herself in a bad situation where nothing better was available, and adapted her preferences to what she could actually get.

In some cases, the concept might not just be used as a justification for questioning a person’s judgements about what is good for her – it might also be used as a justification for undercutting her right to act on those judgements. Even if we are convinced that a person’s judgements about what is good for her are mistaken, most of us will think much of the time that she is still fully entitled to act on them. Think bland food isn’t worth eating? Think being a dentist would be mind-numbingly boring? You likely still object to pushing a person who wants those things to live a different kind of life. But in some cases – especially cases that involve deprivation or injustice, as all of our examples do – many people’s objections to intervening in another person’s life choices start to falter. Think that slavery is a pretty good deal? Think that you should defer to your husband even when you don’t really want to? Again, adaptive preferences help us to explain why interventions might be appropriate in these cases. Our subject only wants those things because she hasn’t had the option of something better – so let’s try to get her choose the better thing instead, even if she says she doesn’t want to.

The concept of adaptive preferences, then, needs to overcome both the presumption that people know what is best for them, and the presumption that they should be able to act in ways that are bad for them if they want to. And it is here that autonomy becomes a natural ally.

While accounts of autonomy, like accounts of adaptive preferences, are wildly different, I take it that Joel Anderson and John Christman are right that the great majority of accounts agree that the concept of autonomy finds its core meaning in the idea of being one’s own person, directed by considerations, desires, conditions, and characteristics that are not simply imposed externally on one, but are part of what can somehow be considered one’s authentic self.

(2005: 3)

And this unifying idea helpfully explains both why we might be suspicious in at least some cases that the objects of a person’s preferences are good for her, and why we might think it’s sometimes appropriate to push someone to live a kind of life she wouldn’t choose for herself.

Consider first the question of first-personal authority over well-being. In at least some cases, most of us will be compelled by the idea that we can be wrong about what is good
Adaptive Preferences as Autonomy Deficits

for us. Perhaps I am convinced that it would be good for me to be a doctor because my science-loving parents have always pushed me in that direction; perhaps I am avoiding my friends because depression has convinced me that nothing is worth doing any more. In each case, many people would still grant that a person’s desires determine what is good for her – they would just deny that these are my true, authentic desires. I want to be a doctor because of pressure from my parents – not because it’s what I really want; when I say I don’t want to see my friends, it’s my depression talking, not me. We can deny my first-personal authority in these isolated cases, because if we pressed deeper, we would see that my reported desires didn’t match my authentic ones. And it is my authentic preferences that ultimately have authority. The concept of adaptive preferences just offers one more mechanism by which our reported preferences fail to be authentic: that is, because our limited options caused us to downgrade our preferences to what we can get, thereby saving us from the frustration of wanting what we can’t have.

Similarly, the permissibility of intervention into a person’s choices varies for many people according to the extent to which those choices represent her authentic desires. When desires very clearly fail to be rooted in a person’s own authentic sense of self, most of us will abandon the intuition that something wrong is done when a person is prevented from acting on her desires. Think of a person whose desires are all implanted by a hypnotist, or a person who makes all of their decisions by flipping a coin. In these cases, there is nothing about the content of the person’s choices that connects in any meaningful way with who she is or what she wants – and so no great wrong seems to be done when her choices are over-ridden. Of course, most of us are not under hypnosis, and no one makes all of their choices by flipping coins. But our authentic desires may conflict with our reported desires in more realistic situations, and in such cases many of us believe that we are permitted – or indeed, even obligated – to cause a person to act on their authentic desires. Think of the smoker who is trying to quit and asks you to stop them from smoking no matter how much they plead. Or think of Mill’s famous example of the man who can be physically prevented from crossing a broken bridge – but only long enough to tell him that the bridge is dangerous (1859). In these cases, it is the authentic preferences to cross the river safely and to quit smoking that have a hold on us as third parties, rather than the inauthentic preferences for a cigarette or for crossing a bridge one only believes to be safe. And the appeal to adaptive preferences is once again supposed to explain why we are permitted to act contrary to the expressed desires of the person who has them: in this case, there is a conflict between what their authentic, autonomous desires would be, and what their inauthentic, adaptive preferences actually are. Forced to choose, we should act on their autonomous desires rather than their adaptive ones.

In appealing to autonomy, then, recognizing adaptive preferences doesn’t require us to jettison either the widely held presumption that persons have first-personal authority over their own good, or the perhaps even more widely held presumption that persons should be free to act in ways that harm themselves. Instead, the exception proves the rule: by treating adaptive preferences as autonomy deficits, we have a unified explanation of why those presumptions hold in the first place, and of precisely why one or both presumptions fail to hold in the case of adaptive preferences.

2. The Easy Cases

For those interested in autonomy (as I assume readers of this handbook are!), it may seem obvious that the category of autonomy must be quite a bit broader than the category of adaptive preferences. But what is obvious in principle is often far less obvious in practice. While it ought to be

273
obvious to all of us that the term “autonomy” is used in an extremely broad range of ways in the literature, confusion and instances of talking past each other occur far too often when one author erroneously assumes that another is using the term in the same way that they are.

So while it might be obvious that not all autonomy deficits are adaptive preferences, fleshing out the substance of that obvious claim is in no way trivial. And indeed, doing so is important for at least two reasons. First, purely conceptually, identifying the particular subset of non-autonomous preferences that adaptive preferences are will be indispensable for offering a defensible autonomy-based account of adaptive preferences. An account that fails to do this will not only be too broad and vague to help us understand the concept with any degree of sophistication; such a failure will also leave such an account unable to respond to objections that might appropriately be leveled at other areas of autonomy, but need not apply to adaptive preferences. Second, at least in political philosophy, adaptive preferences are often identified in order to be intervened upon – and the interventions that are appropriate for adaptive preferences properly understood may be very different from the interventions appropriate for other kinds of autonomy deficits. If we fail to fill in the contours of the distinction between adaptive preferences and autonomy deficits more broadly, we might recommend adaptive preference interventions that would be fully appropriate for some other types of autonomy deficits, but completely inappropriate for adaptive preferences.

As I said above, I won’t aim here to make a case for a particular positive account of adaptive preferences as autonomy deficits. Rather, I’ll just lay out the kinds of boundaries that positive accounts might reasonably draw, and leave arguments in favor of one account or another for other venues. In this section, we’ll consider the kinds of autonomy deficits that should clearly not be understood as adaptive preferences on any plausible account, returning to cases of reasonable disagreement in the final section.

In order to identify the set of autonomy deficits that are obviously unrelated to adaptive preferences, let’s step back for a moment and try to canvas the range of ways in which accounts of autonomy might treat a preference as suspicious or inauthentic. Once we’ve done that, we’ll be in a much better position to determine which subset of non-autonomous preferences adaptive preferences might be. I’ll begin by dividing failures of autonomy into two types: genesis and content. We’ll discuss genesis in this section, reserving content for the section in which we discuss reasonable disagreements between adaptive preference theorists.

One way in which preferences can warrant suspicion is via their genesis – that is, the circumstances under which they were developed and maintained. Given the centrality of authenticity to accounts of autonomy, the ability to withstand self-directed critical scrutiny is a common requirement for the autonomy of preferences. For this kind of test to be useful for our purposes, however, we must recognize that such failures can be of two very different kinds: broad failures of capacity, or more particular failures of exercise. In the first case, a person fails to critically reflect on a preference because she lacks the ability to do so. If a person has an extremely low capacity for critical reflection, then critical reflection cannot do much to authenticate her preferences. In the second case, a person has the capacity for critical reflection but, for some reason – laziness, depression, exhaustion, etc. – fails to exercise it. While these two kinds of failures may offer very different prospects for possible future successes of critical reflection, the immediate outcome for the preference will often be the same. If laziness, for instance, prevents a person from reflecting in any way upon a preference that she has formed, then it should be almost as suspect as the same preference held by a person who almost entirely lacks the ability to critically reflect at all. In both cases, persons are unlikely to have happened to alight upon and so prefer the option that is most authentically theirs.
Lack of capacity and failure to exercise it are both plausibly understood as internal failures of genesis, but genesis must also consider external circumstances. Whether a person critically reflects on her preferences as she develops them affects the degree of suspicion that those preferences warrant, but the circumstances under which she undertakes this reflection are also important. If a preference is developed without even an awareness of alternatives, then critical reflection will do comparatively little to determine whether it is a preference that should be maintained. A preference developed in this way, then, should be treated as much more suspect than one that is upheld in light of various alternatives. For instance, a person’s preference to study mathematics will be much less suspect if she has taken some courses in other subjects and rejected them than if she has only ever tried math.

So far, then, a preference’s autonomy can be affected by the following sorts of failure in genesis: external failures of circumstances or internal failures of capacity, with internal failures of capacity further divided into global lack of capacity and local lack of exercise of that capacity. Taken together, however, internal and external failures of genesis can lead to a third cause of suspicion. Think of a case involving a hypnotist. Imagine that I have the capacity for critical reflection, but that a hypnotist induces in me a preference to study mathematics at university which replaces my previous preference to study philosophy. Should the reason to second-guess this preference be taken to be a failure of internal critical reflection or a failure of conducive external circumstances? There are cases to be made in each direction. Clearly, the failure of critical reasoning is a result of external circumstances. If the hypnotist had not intervened, then the failure would not have occurred. The real problem with the preference, however, is not that it was developed without awareness of alternatives in the external environment, but rather that external circumstances compromised the internal process of critical reflection that would have allowed them to be seriously considered. Such combined instances of internal and external sources of suspicion should be treated separately, since different things will be required to overcome suspicion in such cases. Cases that are only failures of external circumstance will only require the introduction of new alternatives. Cases that are only failures of the exercise of capacity will require, straightforwardly, exercising the capacity. But when external circumstances have been internalized to the extent that they make a person disinclined to critically consider alternatives that become newly available, overcoming suspicion will mean not just becoming exposed to new alternatives, but also becoming willing to see them as alternatives that are genuine and that deserve critical consideration.

Clearly, all of the sources of suspicion that arise from genesis are relevant to the autonomy of a preference. Failures of capacity for critical reflection, failures of external circumstance, and internalization of limited circumstances all rob persons of the tools to exercise effective self-rule; the failure to exercise critical reflection simply leaves these tools unused. But adaptive preference theorists should be unified in holding that not all of these genesis-sourced bases of suspicion warrant a judgement of adaptive preference. The word “adaptive” itself is necessarily attached to the idea of adapting to circumstances, and the kinds of intuitive examples we used to flesh out the concept involved preferences that warranted suspicion because they were formed in accordance with circumstances of social injustice, inequality, or domination. Accordingly, therefore, simple failures of exercise of capacity for critical reflection are not relevant to adaptive preference, since they are not caused by circumstance. Failures of global capacity might be related to social injustice or inequality, as when infants in impoverished households fail to receive the medical care, special education, or nutrition that they need to develop their capacity for critical reflection. However, although their future preferences will likely be affected by the failure to develop this capacity, the suspicion that their preferences warrant will be a result of direct causation by their circumstances,
rather than moderated adaptive response to them. Accordingly, adaptive preferences theorists should also be unified in taking deficiencies in capacity to fall outside of discussions of adaptive preferences proper. Both external and internalized failures of circumstances, on the other hand, will be relevant to discussions of adaptive preferences—the former because it keeps persons from becoming even acquainted with the alternatives that might be better for them, and the latter because it prevents persons from seriously considering those alternatives should they become available.

3. The Hard Cases

Let us now turn to points of reasonable disagreement that might arise between adaptive preference theorists. Perhaps the most significant point of disagreement is whether adaptive preferences should be understood as autonomy deficits at all. However, as I said above, this chapter aims to lay out the various ways in which we might understand adaptive preferences using accounts of autonomy. Accordingly, I leave this primary point of reasonable disagreement to the side, and will focus here only on reasonable disagreements between those already committed to autonomy-based accounts. In the previous section, we saw that adaptive preference theorists should be unified in holding that autonomy deficits are not adaptive preferences if they are due to a global lack of capacity for critical reflection, or if they are a simple failure to exercise that capacity that is in no way related to deficient circumstances or choice sets. Beyond this, however, adaptive preference theorists might reasonably disagree on how narrow a subset of non-autonomous preferences adaptive preferences are. In this section, I’ll consider three significant reasonable disagreements, although I don’t take this discussion to be exhaustive. I’ll begin with the values involved in the preferences, move on to the centrality or importance of the preferences in a person’s life, and end with a discussion of the role that a person’s conscious intentions play in the adaptation of their preferences.

First, let’s consider the role that content might play in counting preferences as adaptive or not, with a focus on the extent to which the object of the preference in question has a certain kind of value. The value of a preference’s object will be most obviously relevant to those committed to perfectionist accounts of well-being. Once you have committed to an account of objective value, a clash between the content of that account and the content of a preference is an obvious reason to treat the preference as liable to suspicion. Indeed, this is the route employed by both Martha Nussbaum (2000) and Serene Khader (2011), two of the most prominent adaptive preference theorists. Those in this camp will label a preference as adaptive when it seems to be bad for a person, unworthy of a human being, flourishing-incompatible, or something else along these lines. And while one need not be a perfectionist in order to count a preference as liable to suspicion on the basis of its objective content, perfectionists will be more likely than non-perfectionists to take the substantive content of a preference to be relevant to its adaptiveness.

Sources of suspicion related to content, however, may also concern subjective or personal value commitments. Here, a preference will be liable to suspicion not because its content conflicts with the content of some purportedly objective list of values, but rather because it exhibits a conflict between a person’s own various subjective value commitments. This sort of conflict is compatible with endorsing an objective account of value, and may be cause for concern even when no preferences that contravene that objective account exist. Any plausible account of objective value will need to be both broad and pluralist, and different elements of such an account may conflict with one another. Even on an objective account of value, then, a person may hold conflicting value commitments that would each be acceptable.
Adaptive Preferences as Autonomy Deficits

on their own. The fact that they conflict, however, should make both preferences liable to suspicion since satisfying one will necessitate frustrating the other—a state of affairs that will almost certainly make a person worse off on at least some important measure than she would be if none of her preferences needed to be frustrated. Note the way, then, in which this element of the content category relates to the genesis category: suspicion on the grounds of subjective content will often be warranted precisely on the grounds that it suggests that a person has failed to critically reflect on her preferences and bring them into line with one another.

Now, either of these categories may be relevant to autonomy. If it is the case that there are objective values that exist independently of human beings, then failing to choose in accordance with them might compromise the quality of your self-rule. Conflicts between subjective or individual values are even more obviously related to autonomy. Insofar as a person has conflicting values, she will always fail to regulate her life in accordance with some of them. Either of these categories may, as well, be relevant to adaptive preferences. As noted above, prominent adaptive preference theorists make objective content a central element of their definitions of adaptiveness, requiring that preferences compromise human flourishing or thickly described capabilities in order to count as adaptive. Subjective value conflicts, as well, may be the effect of adaptation to unjust circumstances. Imagine, for instance, that a woman develops a desire to do groundbreaking scientific research, but also develops a distrust of women as serious scientists because of the sexist norms that surround her. Even if there were nothing inherently wrong with sexism, this woman would be unable to direct her life in accordance with all of her central commitments and beliefs.

There is an argument to be made, then, that both objective and subjective value conflicts may lead to cases of adaptive preference. Adaptive preference theorists, however, can reasonably disagree about whether both types should be included—and the theory of autonomy on which they rely provides one ground on which they might disagree. Autonomy theorists committed to strong substantive accounts of autonomy may be more inclined to adopt an account of adaptive preferences that makes objectively harmful content a necessary condition of adaptiveness, since they have already endorsed substantive content as a requirement of their accounts of autonomy (i.e. Benson 1991 and Stoljar 2000). Content-neutral theorists, on the other hand, may be less willing to require particular content for adaptiveness. Since content-neutral theorists so often reject substantive requirements for autonomy on the grounds that it imposes a value set on those who might reasonably reject it (i.e. Friedman 2003 and Christman 2009), they may also be inclined to reject substantive requirements for adaptiveness on the grounds that people may reasonably disagree about what is compatible with human flourishing.

Next, let’s consider a second but orthogonal issue of content: what is the importance of the preferences in question? While the answer to this question may not be determined by the content of one’s account of autonomy, as the last question plausibly was, adaptive preference theorists will nevertheless need to take a stand on this issue. What is at stake here is how expansive we want the category of adaptive preferences to be. Some cases that fit our vague, structural definition of adaptive preferences—remember, preferences you endorse as a result of bad or limiting circumstances, that you would not have developed or endorsed in more expansive circumstances—are clearly of great importance for the people who hold them. The enslaved life that our Happy Slave endorses, for instance, is bad for her in extreme and all-encompassing ways. But you might also have much more trivial preferences that fit our vague criteria. Think of the person who has developed a taste for extremely bland food, after never having been exposed to diverse or interesting flavors. While this kind of preference has the same form as the preference of the Happy Slave, it just matters less: it involves only
Rosa Terlazzo

one narrow area of her life, and it doesn’t seem to make her that badly off even in that very narrow area of her life.

Adaptive preference theorists, then, will need to decide whether or not they want to include trivial preferences in their account alongside the consequential ones. Although, as I mentioned, your account of autonomy might not help you to make this decision, your purpose for adopting an account of adaptive preferences in the first place might. If your purpose is primarily political – that is, if it is primarily focused on identifying adaptive preferences that warrant intervention, and/or on justifying that intervention – then you will likely want to exclude trivial preferences from your account. Even if your preference for bland food makes your life worse, political intervention aimed at changing this kind of preference will strike many of us as highly objectionable. If, on the other hand, your purpose is primarily descriptive – that is, if you are primarily focused on understanding the impact that adaptation to circumstances should have on our judgments about well-being – then you will likely want to include the trivial preferences. Since many trivial features of our lives can add up to much better or worse lives than we might otherwise have, we should aim to understand the well-being contributions of even trivial adaptive preferences.

Note that political projects will likely also want to make a further distinction that descriptive projects may not need to make: namely, between deeply consequential preferences that are politically objectionable and deeply consequential preferences that may matter significantly for well-being without being worthy of political intervention. Consider someone who we’ll call the Reluctant Doctor, who became a doctor due to generations of family pressure and now lives a miserable life that they nevertheless endorse as the best kind of life for them. While the Reluctant Doctor’s life might plausibly be just as bad for her as the Deferential Wife’s, the latter seems to be appropriate for political intervention in a way that the former does not. In addition to deciding on the purpose of one’s account of adaptive preferences, then, any adaptive preference theorist will have to adopt a supporting theory that explains either the difference between the Deferential Wife and the Reluctant Doctor, the difference between the Reluctant Doctor and the Bland Eater, or both.

Third and finally, let’s return to genesis, to consider the role that awareness and intentionality in a preference’s development should play in determining its adaptiveness. Our preferences can be modified according to circumstance both in intentional ways and in ways that occur without our conscious awareness, and adaptive preference theorists will have to decide whether or not both kinds of modification will be included in their account. Jon Elster (1983), for instance, explicitly distinguishes adaptive preference from character planning, and Ben Colburn (2011) limits adaptive preferences to those that are developed “behind our backs.” Part of the point here is to explain how one might respond to limitations in ways that are both authentic and prudentially wise. We all face limitations, and many of us will face limitations that put some of the lives that we most want to live out of our reach. Think of all of the children who aspire to be professional athletes, but just aren’t strong enough or fast enough or tall enough to make it happen. These children would be wise to intentionally modify their preferences so that they desire future lives and careers that are a real possibility for them, and some adaptive preferences theorists might accordingly want to place this kind of modified preference outside of the sphere of adaptiveness, in order to signal that such preferences are not criticizable in the way that others are.

In this case, your account of autonomy might again influence the route you take. If your account of autonomy is primarily focused on the existence or exercise of the capacity for critical reflection, then you might be inclined to count only unconsciously modified preferences as adaptive. After all, if your focus is on the extent to which you have critically engaged with
Adaptive Preferences as Autonomy Deficits

and endorsed your preferences, then this kind of intentional modification could seem to be autonomy-enhancing rather than autonomy-limiting. If, however, you have an account of autonomy that is relational or that otherwise places great store on the ways in which external circumstances can constrain our authentic selves, you might want to count as least some instances of intentional preference modification as adaptive. Think, for instance, of the gay man in 1950’s America, who intentionally develops a preference for a companionate marriage to a woman in order to avoid the punishments of a deeply homophobic society. Even if he has consciously adopted this route, many people may see this as a case in which we should recognize external constraints as limits on autonomy – in large part because they think we should see it as a case in which his consciously formed preference is adaptive.

Conclusion

Understanding adaptive preferences as autonomy deficits is both plausible and attractive. Deciding that we should understand them in this way, however, cannot be the end of the story. Instead, adaptive preferences theorists who want to take this route must be clear about which subset of non-autonomous preferences adaptive preferences are. Doing so will enhance our understanding of adaptive preferences for obvious reasons, but it should also enhance our understanding of autonomy. By taking a stand on the points of reasonable disagreement between adaptive preference theorists, we may also end up with a better sense of what we want the contours of our more general accounts of autonomy to be, and what we want the judgment of an autonomy deficit to license.

Notes

1 I follow the convention in the literature of using “preference” in the broad sense that includes things like values and desires, rather than in the narrow sense of the three-place predicate where a preference always and only involves agent A’s relative judgments about particular objects B and C.
2 For arguments against understanding adaptive preferences as autonomy deficits and responses to those arguments, see, respectively, Khader (2011) and Terlazzo (2016).
3 There are real and serious concerns that such reasoning both harms and disrespects the people it aims to help. I share many of these concerns, but lack the space to address them here. For more on these issues, see Jaggar (2006), Narayan (2002), and Khader (2011).
4 For an extended argument against understanding adaptive preferences as autonomy deficits, see Khader (2011).

References


