ASKING MESSY QUESTIONS ABOUT ADAPTIVE PREFERENCES

Rosa Terlazzo
University of Rochester
Forthcoming in *Australasian Journal of Philosophy*

Careful, precise, and important work has been done on a range of questions in the adaptive preferences literature. In this paper, I make the case that we should be doing more of the messy work at the intersections of these separate questions. Otherwise, I argue, we miss out on some of the most important insights and puzzles about adaptive preferences. In the paper, I consider the following questions in light of each other: 1) In what modes and orientations should we consider adaptive preferences? 2) Are adaptive preferences actually bad for their holders? 3) Should we want to try to change them? and 4) If we should try to change them, who is the 'we' who should be trying to make the change? Through discussion of these questions, I conclude that people with adaptive preferences may rightly recognize that the satisfaction of their adaptive preferences benefits them, and may reasonably therefore not want to try to change their adaptive preferences. However, I argue that this does not settle the question of whether they should all things considered aim to change their adaptive preferences. Instead, I argue that people with adaptive preferences must be seen both as moral patients with prudential interests at stake, and as moral agents who bear obligations to others. But we only have access to this insight if we simultaneously consider the questions of what adaptive preferences are, whether satisfying them is bad for their holders, whether we should want to change them, and who should be doing the changing.
Keywords: adaptive preference; agency; obligation; well-being

The literature on adaptive preferences is over 40 years old. It contains rich debates on the nature of adaptive preference (Khader 2011, 2012; Nussbaum 2000), whether adaptive preferences are bad for their holders (Dorsey 2017; Terlazzo 2015b), whether they deserve our respect (Jaggar 2006; Khader 2011), whether and how we should aim to change them (Nagar and Raju 2003; Fuller 2011), and whether we should attribute them to people at all (Baber 2007) – to name just a few. Each of these literatures is careful, precise, sophisticated, and important. But in that large literature, I am increasingly struck by what is lacking: work at the messy places where these questions intersect. And I am increasingly convinced that we miss out on some of the most important questions about adaptive preferences when we fail to consider the extant questions in light of each other. In this paper I won't aim to definitively answer any of the questions in the literature. Instead, I'll aim to show what we miss if we do not consider them together – that is, what we miss when we fail to get messy.

In this paper, I'll focus on four different questions. First, how narrowly should we understand adaptive preferences? Second, are adaptive preferences in fact bad for those who hold them? Third, where adaptive preferences exist, should we try to change them? And finally, if we should try to change them, who is the 'we' who should be doing the changing? The argument proceeds as follows. In Section I, I explain the broad way in which I'll be understanding adaptive preferences in the paper. In Section II, I argue that while much of the literature on adaptive preference change focuses on interventions by governmental or non-governmental organizations, even these kinds of interventions involve an important role for individuals as agents of adaptive preference change. In Section III, I show how the interests of individuals in adaptative preference change can come apart from the interests of organizations, as well as from the interests of other members of their group over time. In Section IV, I introduce four different ways in which individuals might genuinely benefit from the satisfaction of their adaptive preferences. Finally, in Section V, I return to the questions of whether we should want to change adaptive preferences, and if so, who should be doing the changing. I argue that most approaches to the problem of adaptive preferences consider individuals with adaptive preferences only as bearers of interests and not as bearers of obligations, while a satisfactory answer to the question of whether we should want to try to change adaptive preferences must consider them as both. Ultimately, I conclude that bearers of adaptive preferences might rightly recognize that they are made best off for their own sake by satisfying their adaptive preferences – but that they might nevertheless have reasons or even obligations to try to change those adaptive preferences for the sake of future generations. And throughout I show how we miss this important insight if we fail to consider our four questions together.

1. What are adaptive preferences?

The term 'adaptive preference' originally comes from Jon Elster, who used it to describe the phenomenon of sour grapes: the fox in Aesop's fable wants the grapes hanging above him, but upon discovering that they are out of reach, decides that they must have been sour and that he therefore doesn't want them anyway (Elster 1983). The example shows us the problem for social welfare calculations clearly. If a person's desire not to have a thing is a reaction to the fact that she *cannot* have it, then her preference not to have it seems to do a poor job of capturing her welfare. Social choice

functions that take her preference not to have the metaphorical grapes as given will frustrate her welfare rather than advancing it.

As the literature on adaptive preferences expanded – initially through Amartya Sen and Martha Nussbaum, and later more broadly into feminist philosophy, development ethics, and the literature on autonomy and well-being – the motivation for thinking about adaptive preferences has remained largely the same. If we defer to people's own preferences, and people's preferences can be adapted to accept or even endorse their own domination and deprivation, then deferring to their preferences will only reinforce that domination and deprivation.

In virtually all other ways, however, the literature has fractured and Elster's original characterization and focus have fallen away. Consider an utterly non-exhaustive list of current debates about adaptive preferences: Should they include deviations from Elster's original formulation, such as life-long habituation (Nussbaum 2000) or some forms of deliberate character planning (Terlazzo 2015a)? Should we build substantive normative commitments into our accounts of adaptive preferences, such as theories of basic flourishing (Khader 2011; Nussbaum 2000)? Should we be sceptical of attributions of adaptive preferences (Baber 2007; Narayan 2002), or think that they are disrespectful (Jaggar 2006)? How should we respond in practice to preferences that we suspect are adaptive (Ackerly 2008; Fuller 2011; Khader 2011)?

Clearly, no one could settle all these debates in one paper. Mentioning them, however, will help in two ways to frame the way that I will understand adaptive preferences here. Less controversially, I will share the motivation that animates the debate about adaptive preferences in development ethics and political and feminist philosophy,[1] and concern myself only with preferences that are in some interesting sense connected to one's own oppression, deprivation, or domination.

More controversially, however, my understanding of what counts as an adaptive preference will be broad along two dimensions. First, consider what you might call their *mode*. Some people, including Elster and Nussbaum, understand adaptive preference as a kind of unreflective drive that can be contrasted with intentional character planning (Elster 1983) or informed desire (Nussbaum 2000). I will reject this narrowing of the concept, also considering preferences that are more intentional and

[1] Note that some other related literatures – for instance, in well-being (Dorsey 2017) and autonomy (Colburn 2011) – legitimately question this focus rather than treating it as a starting point.

informed (as long as they still respond to or maintain one's own domination or deprivation in the right kind of way). I'll return to why below.

Second, I'll understand adaptive preferences broadly in terms of what you might call the subject's *orientation* towards the preference's object.[2] Consider three different orientations you might have towards an object you pursue. First, you might take the object of your preference to be a bearer of value worth endorsing or desiring for its own sake. It is in this sense that many people prefer living in a walkable neighbourhood or being a member of a religious congregation. Second, you might engage in a more limited and comparative form of valuing, where you take one object to be *more* worthy of endorsement or desire than another, but where this comparative endorsement tells me little about your more general preference ordering. It is in this sense that you might prefer the friends that you make in the small college town you have moved to. And third, your pursuit of an object might merely reflect the fact that you take it to be the most choice-worthy option from the set available – that is, it might be simply what an economist would call a revealed preference. It is in this way that we might describe your choice to work at McDonald's rather than in an Amazon warehouse, or to pay a fine rather than doing community service. Note that while the second and third uses of the term are both comparative, they differ in that, in the third use, choosing an object need not imply that you value it for its own sake at all. In the friendship case, if you attached no value to friendship, you could simply choose a solitary life. But in the revealed preference sense, you might hate both job possibilities but still choose one in order to avoid the worse alternative of being unable to pay your bills.[3]

As I said, including all these orientations and modes as cases of adaptive preference is controversial. Some, like Elster or Nussbaum, might say that preferences in the reflective mode are not adaptive at all, since they are consciously reflected upon and chosen. They will say that preferences with the same content deserve respect when they are reflective, but not when they are adaptive. And others, especially in the feminist and development ethics literature, will say that adaptive preferences of the revealed kind (or perhaps even of the comparative valuing kind) are not really *preferences* at all – instead, they are attempts to get by within a system that one does *not* in any way prefer, and so do not raise the tricky

[2] The stricter sense of preference as a three-place predicate where *a* prefers *b* to *c* is more generally implicitly rejected in the literature in favor of an understanding of preference that includes broader kinds of non-comparative endorsement. I follow this convention here.

[3] Note that for the economist, revealed preferences are meant to describe all of our actions in the world, including those that involve genuine valuing. For the sake of conceptual clarity, I will use the term to refer only to the third type of orientation, which involves choice without value – although I of course recognize that our actual choices are often based in our values.

problem of how to respond to a person's preference for their own oppression (i.e. Baber 2007). Here there is no problem about whether and how to respect their preferences for oppression, because what they would in fact *prefer* is to not be in the oppressive situation in the first place. In short, the objection in both cases goes, some modes and orientations of preference simply raise problems that other modes and orientations do not.

I, however, will be considering adaptive preferences in both modes and all three orientations. I recognize that adaptive preferences in different modes and orientations raise problems of respect and response in different ways. A considered, informed, and enthusiastically-endorsed preference for a norm that marginalizes you ought to be taken more seriously than an ambivalent and unconsciously held drive. An unconscious tendency might warrant an invitation to reflect in light of alternatives, while a clearly-considered and -expressed preference might require more directly grappling with the question of whether to facilitate or frustrate it. But I include them all because of what we lose when we leave any out. And that case will be best made by showing in each of the following sections what is lost when we do so.

One final preliminary point before we turn to our three main questions. Questions of whether adaptive preferences should be changed are at the heart of this paper – and readers who come to the debate via Elster or Sen might object to this way of framing the issue.[4] After all, Sen and Elster are both concerned with social choice – they are interested in whether adaptive preferences should be *incorporated* into social choice functions, not whether they should be *changed*.

Strictly, this is true. But if we take a step back, there are two important reasons to think that we should still take questions about whether to change adaptive preferences to be at the centre of our debate. First, social choice functions are not merely objects of academic interest. Instead, they are *tools* that help us to take seriously all interests in society when deciding what our social policies should be and what we ultimately want our society to look like. If adaptive preferences are taken as given by social choice functions, then (all other things being equal) the policies that our functions recommend will aim to satisfy those preferences (or at least not undermine them). But if adaptive preferences are not included on the grounds of their adaptiveness, then the resulting policies will not aim to satisfy or maintain them. This is not yet to say that Elster or Sen's work has the aim of changing adaptive

[4] I am grateful to an anonymous reviewer for raising this worry.

preferences – but it should be enough to show us that their work has significant *implications* for whether adaptive preferences will be changed.

To tie us more tightly to the aim of *changing* adaptive preferences, consider the kinds of examples that Sen in particular uses to motivate his interest in adaptive preferences. Sen asks us to consider 'The hopeless beggar, the precarious landless labourer, the dominated housewife, the hardened unemployed or the over-exhausted coolie', all of whom take pleasure in small mercies and suppress their suffering for the sake of survival (1987:45). The idea is not that preferences of this sort should be excluded from social choice functions for reasons that remain neutral on their content (as, for instance, in the case of other-regarding preferences that ought to be excluded because they involve double-counting). Rather, Sen offers us these examples because they make a claim on us morally: these people's interests matter, and we should want a society in which people can both form and satisfy preferences that are better for them. And indeed, the vast majority of the literature on adaptive preferences that follows on from Sen and Elster is focused precisely on how, when, why, and whether we should engage in interventions aimed at getting people to form and satisfy different, non-adaptive preferences (for example, Fuller 2011; Khader 2011; Nagar and Raju 2003; Nussbaum 2000). Thus, even if changing adaptive preferences was not the initial focus of Elster and Sen's work, I take the question to be central enough to the literature to require our attention.

With this preliminary discussion out of the way, let's turn to our remaining questions.

2. Who are the relevant change agents?

We turn now to the question of agency: if we aim to change adaptive preferences, who should the agents of change be? When the problem is framed as one of changing society in order to foster different and better preferences, the most obvious agent of change will be either some kind of government body that will pass a relevant law or policy (such as outlawing female genital cutting), or some kind of non-governmental organization that will undertake outreach or incentive programs aimed at changing behaviour (such as providing micro-loans to women but not to men, or forming

consciousness-raising groups). Without this sort of coordinated, wide-scale action, society is unlikely to change.

But we can also think of individuals as the relevant change-makers in cases of adaptive preferences. In the distinct but related literature on the obligations of victims to resist oppression, the focus is on what victims themselves must do in order to respond to oppressive systems with integrity (i.e. Hay 2011; Khader 2021; Terlazzo 2020). As Ann Cudd notes, oppressive systems are so intransigent in part because they so often function precisely by making victims complicit in their own oppression (2006: 146ff). In these cases, systems are maintained by norms with three features. First, the norm is maintained through the compliance of individual members of the group in question. Second, the group in question is made worse off by the existence of the norm. And third, given the existence of the norm, individuals who refuse to comply with the norm are made worse off than those who acquiesce. Think here of feminine beauty norms involving make-up, clothing, hair removal, etc. If no individual women complied with them, those norms would no longer exist. They also make women as a group worse off, by requiring women to invest significant time and money in enacting them, and harshly punishing women who through age, disability, body size, or some other feature, are unable to meet their standards. And finally, women who could comply with them but refuse to do so are also made worse off: women who do not meet them are judged to be less professional and are punished in the workplace; the internalization of these beauty norms mean that women who do not meet them face a more limited range of potential romantic partners; women who do not meet them can encounter harassment or abuse in public, as many fat women or butch women can attest; the list goes on.

In these sorts of cases, the oppression involved is generally not a direct result of a government policy that could be implemented or repealed. Instead, because the norms are in large part informally maintained, and because their maintenance depends upon the compliance of the oppressed population, what is required for the norm to be undermined is the large-scale refusal of individuals to comply. While this literature has focused on obligations of victims to resist oppression, the implications for adaptive preferences are clear. In many cases, the circumstances that marginalize people and lead to the development of adaptive preferences are comprised at least as much by informal norms as by formal laws or policies. Many women enjoy make-up, high heels, and fashion, and many say that they find themselves less beautiful without make-up, or report distress and disappointment when they weigh more than restrictive beauty norms dictate. Women have internalized these norms because they are so pervasive, and it is hard to see how they would have formed the same preferences

if women did not so widely (try to) conform to them. So in many cases, while governments might pass tangentially-related laws, or NGOs might engage in consciousness-raising activities, changing the norms that lead to adaptive preferences will often require action on the part of many individuals who face the choice of whether and how to comply with the oppressive norms that apply to them. And even if effective governmental policies or non-governmental programs are undertaken, their effectiveness will often be a function precisely of individuals choosing to comply with them.[5]

So when we think about agents of change capable of undermining adaptive preferences, we should be thinking at least as much of individuals as of governmental or non-governmental bodies.

3. Should we try to change adaptive preferences?

Recognizing that there are at least two sets of potential change agents involved complicates the question of whether we should want to try to change adaptive preferences, because the interests of these two groups may well come apart.

First and perhaps most obviously, the government or non-governmental group might have an interest in changing norms and associated preferences that is not shared by individuals being asked to undertake the change. Think here of our own profession of philosophy. The discipline would plausibly be much better off if it were more diverse along the lines of race, gender, dis/ability, and socioeconomic class. This might be so for purely moral reasons, because diversity is valuable and exclusion is bad. Or perhaps it is so for more instrumental or epistemic reasons – because its narrow demographics cause it to miss out on a diversity of ideas and arguments, or because they lead to negative perceptions from other areas of universities.

[5] Note that whether governments are required or even permitted to intervene on self-regarding preferences that marginalize some group in society is an independent question. While I think that there are some good reasons to think that these kinds of interventions can indeed be justified in liberal democracies (see Gheaus 2023 and Schouten 2019), I take it that the fact that these choices marginalize vulnerable people is at minimum enough of a reason to want them to be different, and that this reason is one that at least non-governmental organizations may permissibly act on.

The change, however, can't be done by fiat. 'The discipline' – whatever that entity is – might be able to do some things to encourage more diverse practitioners to pursue and remain in careers in philosophy. But ultimately, philosophy as a discipline will only become more diverse if individuals from more diverse backgrounds become interested in professional jobs in philosophy, are hired into those jobs, and remain and flourish in those jobs once they have them. And as we've seen, oppressive structures are often maintained by penalizing those who refuse to or are unable to comply with dominant norms – which in this case means the individuals aiming to diversify the profession through their presence and/or attempts to use more diverse methods. While the discipline might have interests in changing the profession by bringing in more diverse members, insofar as the barriers to entry unique to those potential members remain, their own individual interests might lie in preferring some other career – whether in an academic discipline that is already more diverse, or outside of academia all together.

Secondly, the answers about whether we should want to do the work of changing adaptive preferences might also come apart within the affected group over time. Here think of potential diverse practitioners of philosophy in this generation versus the next one. Many women, people of colour, disabled people, and people with low socio-economic status backgrounds have fallen in love with philosophy – and many of them have left because they have found it to be an inhospitable or precarious place to be. These people surely wish that the hard work of changing academic philosophy had already been done by an earlier generation, and those in future generations would certainly benefit from the work being done now. But for potential diverse practitioners *actually* making their choices now, the barriers may simply be too high. They may very reasonably not want to spend their lives doing the work of changing philosophy.

At this point, we begin to see why we must consider adaptive preferences broadly in terms of both mode and orientation. In terms of orientation, if we understand adaptive preferences only as the valuing of an object for its own sake, then our potential diverse practitioners do not seem to have an adaptive preference for philosophy to remain a homogenous, narrowly populated discipline. It seems that they would prefer in the deepest sense that they faced no barriers to entry, and could pursue their love of philosophy considering only preference in the sense of weighing the real value they see in philosophy against the real comparative value they see in other potential lives. In this sense, one major feature that undermines adaptive preferences as objects of our respect don't seem to be present: these people don't seem to prefer their own oppression, but instead to prefer to exit the profession that

marginalizes them. So questions of respect are not raised when we consider whether to treat the preferences as authentic or inauthentic. But it seems to me that the preferences *are* nevertheless undermined in several of the other important ways that generally come with adaptiveness: they are preferences for abandoning an otherwise-preferred option that are conditioned by oppressive structures and that in turn strengthen those oppressive structures. And these other features raise a very different question of respect: do these preferences demand respect in the sense that they deserve to be facilitated in our social policies, or is their demand on us in some sense undermined by their oppressive genesis and the role they play in reinforcing oppression?

Now consider mode. If, like Elster or Nussbaum, you think that adaptive preferences are only unconscious drives, then potential diverse practitioners who have reflected on their options and decided to leave will not count as having adaptive preferences. But notice that the case is still highly relevant for those interested in whether to change adaptive preferences. Consider all the women and people of colour who never fall in love with philosophy in the first place because of some inchoate sense that they are not welcome there. These seem to be genuinely adaptive preferences on any account – but the possibility of changing them still seems to hang on the choices of individuals who might legitimately not want to do the work of changing the discipline. And if we define adaptive preferences narrowly in terms of either mode or orientation, we miss out on these important connections.

To understand better why individuals with adaptive preferences might not want to do the work of changing them, however, we should now turn to our third primary question: namely, are adaptive preferences actually bad for their holders?

4. Are adaptive preferences bad for those who have them?

As I said above, I am interested in preferences that are connected to the holder's own oppression, marginalization, or deprivation. And we might think that so defined, adaptive preferences are necessarily bad for those who have them. But this would be too fast. Even in the case of these

worrisome adaptive preferences, the objects of preferences in question might end up being good for their holders in important ways. I'll canvass four ways in which this might be so, and then return once again to why these varieties of preference demand us to consider adaptive preferences broadly in terms of both orientation and mode.

4.1 Benefitting bundles

First, the object of an adaptive preference might benefit you because it is inextricably tied up with something else you value, and you can only get the otherwise valued thing by also getting the oppressive elements of the adaptive preference. Uma Narayan argues that many adaptive preferences are in fact of this type, focusing on women in non-western countries who embrace elements of patriarchal cultures for the genuine value they can bring to their lives, while also regretting the difficulties and harms of the same practices (2002). She discusses the case of Sufi Pirzada women who recognize both the benefits and drawbacks of full veiling practices. While they recognize that their burqas are heavy, constricting, hot, and in one woman's words make her feel 'like a water buffalo', they also value them intrinsically as symbols of their religious identities, and value instrumentally both the respect that they are afforded given this sign of their identity as highly observant women, and the ability to move about more freely in the world due to the anonymity that the burqa provides. These women might well prefer it if they could separate the beneficial and costly elements of wearing the burqa, availing themselves only of the benefits – but as things stand, the beneficial elements of the burqa are not available to them without the costly elements.[6]

4.2 The costs of effort

Another way in which the satisfaction of an adaptive preference might benefit you involves the cost required to achieve a non-adaptive alternative. In the previous case, it was precisely the same feature (or set of features) of an object that rendered a preference for it both adaptive and beneficial, and an

[6] We should not, however, think that this phenomenon of bundling occurs only in 'non-western' cultures. Consider the trans woman who wears make-up and high heels. This woman might share all the objections to feminine beauty norms that we recognized before, including the cost in time and money, and the punishments experienced by women who cannot meet them. Yet as a trans woman, complying with them might allow her to pass for cisgender, thereby protecting her from transphobic violence or preventing strangers from viewing her as an imposter.

alternative that provided the benefits without the costs was possible but not actual. In the current case, there *does* exist an option that provides the benefits without the costs, but that option is itself costly to access – and the significance of the costs of access might count in favour of the satisfaction of the adaptive preference being all things considered better for the person in question. Think of the woman whose partner does little childcare or housework, and who wants this to be different. She might push the point with him, demanding that he do his share, but in many relationships this kind of demand leads to conflict in the relationship, weaponized incompetence, or resentment from both partners. Depending on the degree of resentment and conflict and the length of conflict period before a result of burden-sharing is reached, the woman might plausibly be made better off by forming and satisfying the adaptive preference to do most of the childcare and housework herself.

The possibility that satisfying the adaptive preference might make one best off becomes even more plausible if we consider the possibility that the more desirable alternative, while available, might not ever be achieved. Consider here the woman who decides not to enter into a relationship with a man who does not pull his weight in the first place. There *are* actually men out there who she could have a genuinely egalitarian relationship with, but unfortunately they are rare – even many of the men who are gender egalitarian in principle fail to live up to their standards in practice. If she holds out for one who does, that wait may pay off – but it also may not. And the cost of going without romantic love is a very great one for those who want romantic love in their lives. She might all things considered be made better off by forming and satisfying the preference for a less egalitarian relationship, thereby expanding her set of possible partners. If she wants biological children with a partner, then the cost is even greater, and her chances of finding a truly egalitarian partner in time become smaller.

4.3 Commitment

A third way in which an adaptive preference's satisfaction might benefit you involves the value of commitment. There is a value that comes with being at peace with a place or state of affairs despite its flaws, or loving and accepting another person without wanting them to be different. But places, states of affairs, and people are all complex, and when you live in a society characterized by significant injustice as we do, it will be hard to find any person or place that does not in some way instantiate or reinforce injustice. So coming to accept or love even generally very good people or states of affairs will likely involve coming to terms with features of them that you could prefer adaptively if at all. One

reasonable response to these sorts of traits is to reject them, and to love the person while remaining critical of these narrower features – and indeed, sometimes this kind of response may even be morally required (Emerick 2016). But purely prudentially, one could reasonably want to avoid the tension and distance that comes with this kind of response, and instead prefer to embrace the person, place, or community as they are, warts and all. Without those warts, I would not be dealing with the same person – and *this* is the person I love.

4.4 Transformation

A fourth way in which a person might benefit from an adaptive preference involves the incorporation of the object of the adaptive preference itself into her life (Terlazzo 2018). The bundled case concerned the inability to disentangle the harms of the object of an adaptive preference from the very real benefits that same object brought. The commitment case involved the value of a unified stance of acceptance or love for a person, place, or state of affairs that involved an adaptive object as one of its many constituent parts. This case involves a more direct attachment to the adaptive object itself, where an object that was once harmful and adaptively preferred becomes incorporated into your identity in a way that transforms its value.

It may seem surprising that an adaptive object could come to benefit us in this wholehearted way, but consider this: we rarely talk about adaptive preferences in the most extreme cases of oppression, such as slavery or genocide. Those are cases of clear domination and injustice that don't admit of other explanations according to which those we take to be victims of injustice might plausibly benefit. We are much more likely to talk about adaptive preferences in cases in which a plausible (although seemingly ultimately unsuccessful) explanation is available for how the system might benefit those we think it harms. We talk about women who believe in the complementarity of the sexes because it provides them with both protection and celebration of their own unique strengths, about poor people who support capitalism because they want their merit and hard work to be rewarded in proportion, about black youth who aspire to the great wealth and success that can come from being an elite athlete or musical superstar.

What we must recognize is that even if these systems are oppressive, they are also complex, and many of the values that they appeal to *are* actually values. Finding a space that supports your own unique

talents really is a wonderful thing, as is the experience of success well-earned, or great athletic or musical achievement. And there are, as well, very many kinds of good lives. Those of us growing up in oppressive systems, then, might initially choose suboptimal options that turn into genuinely good lives over time. We might choose a course of life or set of values that is available to us within an oppressive system, and that would be worse for us than another we could have chosen had we had a more expansive option set. But over time each of us builds a life, and the values, people, commitments, and activities that we build our lives around often become inextricably tied up with who we are. So a woman's commitment to a patriarchal religion, or a man's commitment to emotional stoicism, though initially the product of their oppressive circumstances, might over time become such a central part of who they are and what they value that lives of that type become the genuinely best available to them.

4.5 Why treat these cases together?

At this point, we've seen several ways in which the satisfaction of oppressive or marginalizing preferences can meaningfully benefit their holders. But do all of these fall under the umbrella of adaptive preferences? Why treat them together, given the great differences between them?

Let's begin by thinking about these cases of benefit in terms of orientation. Even if we are only interested in adaptive preferences that involve wholehearted valuing of their objects, at minimum cases of commitment and transformation really do seem to count as both genuine preferences and as having objects that really do benefit their holders. The valuing in the two cases is different – with transformation involving valuing the initially objectionable element directly, and commitment involving a valuing of the whole object that extends to valuing of all of its unique parts – but in both cases we are dealing with sincere cases of valuing that both seem to benefit the person and that seem to arise from and favour the subject's own oppression, marginalization, or deprivation. And this collection of features pulls us in different directions when it comes to questions of respect. It's not clear how we should judge either their authenticity or what kind of demand for facilitation they make.

When it comes to mode, these cases of benefit all have the potential to be relevantly adaptive. You might achieve transformation either through a process of unconscious drives or through a program of intentional character planning. If you are pursuing a suboptimal option because of the costs of effort of pursuing something better, you might do so either in a subconscious way or as the result of

deliberation informed by the alternatives. Many of all of our preferences are uninformed and unconsidered, and if we are steeped in an oppressive culture, then we should not be surprised that some of these will be for oppressive objects.

Plausibly in isolation preferences with the same content that occur in different modes or orientations require different responses. But it is worth considering all these types of benefit together because of the great difficulty of *telling* what a preference's mode or orientation is. Consider judgements made from the outside about the kinds of cases often addressed in the literature on adaptive preferences. Does the woman who encourages her daughter to undergo clitoridectomy (Khader 2011: 78ff) really endorse the practice – or is she reluctantly doing it to ensure that her daughter marries well? Is the woman who feeds her husband and male children first even though there is not enough left over for her (Khader 2011: 78ff) doing it simply because it is what has always been done, or as a conscious strategy to avoid abuse? We cannot speak to every person suspected of having an adaptive preference, and even if we could, marginalized people often have rational protective reasons not to express criticism of the systems that marginalize them.

And it is often just as hard to know from the inside how to understand our own preferences. While I have treated these modes and orientations as distinct, in practice the lines are often much blurrier. We are often not entirely transparent to ourselves, and in many cases – especially absent the kind of intense critical reflection that is difficult and costly to engage in in any case – we will be unsure whether the choices we make are in fact the result of deep valuing, comparative valuing, or simply an instance of revealed preference. Indeed, we will often not reflect on our preferences at all. This is an especially important point in cases of adaptive preferences, since acting in ways that disadvantage or marginalize ourselves can create painful cognitive dissonance that we have very real incentives to soothe. One plausible way of doing this is to believe, at least at some level, that our choices do express and align with our values, even if this may not be the case upon scrutiny – or simply to avoid that scrutiny in the first place.

If meaningfully adaptive preferences can benefit us in different modes and in different orientations, and we often will not be able to tell in practice what the modes or orientations of those preferences are, then we had better consider all the possibilities together. If we instead confine ourselves to one 'pure' category of adaptive preference, then our judgements about the respect that adaptive preferences deserve will be too narrow to guide us in another of the central questions that this paper

is concerned with: that is, the desirability of efforts to change adaptive preferences in the world. It is to this question that we now return.

5. The desirability and agents of change, revisited

Once we better understand both the potential benefits of adaptive preferences and the difficulty of separating adaptive preferences in different modes and orientations, it becomes even harder to see whether we should want to change them. We have seen a range of ways in which satisfying adaptive preferences might meaningfully benefit individuals, and this gives them good reasons not to want to change their own preferences – either in the sense of the actions they pursue, or in the sense of their satisfaction with oppressive circumstances that they must act in light of. Choosing to satisfy their preferences gets them some real benefits, and being happy with their state of affairs is much more pleasant than resenting it. Similarly, governments and NGOs should recognize both the real benefits that people can achieve from satisfying adaptive preferences, and their own limited ability to change oppressive circumstances wholesale. If governments and NGOS have the aim of benefitting people, then the fact that people can meaningfully benefit themselves by satisfying adaptive preferences counts against governments and NGOs trying to change either those preferences or the associated behaviours.

But this does not yet settle the question of whether we should want to change adaptive preferences. Remember the future potential generations of more diverse philosophers, who would benefit so much if the hard work of changing professional philosophy were being done now. Our potential diverse practitioners will have a corollary in any case in which an oppressive norm harms a group of people in the aggregate, but punishes individuals for failing to comply with it. Future generations will always be made better off if we do the work of undermining that norm now, so that they may act unconstrained by it when their time comes. Surely their interests count too. While their future interests may not settle the question of whether we should try to change adaptive preferences now, they must at least be part of the equation. And this once again gives us a reason to treat adaptive preferences in their various orientations and modes together. Whether adaptive preferences are informed or inchoate, and whether they involve genuine valuing or reluctant trade-offs, they create the same forward-looking questions about what to *do*: Should they be facilitated because of the way in which

they benefit their subjects now? Or should they be undermined in order to undermine the circumstances that will put future generations in a position to develop adaptive preferences?

To think more carefully about how these sets of interests should be incorporated, let us return to the question of who our agents of change actually are. When I first addressed this question, I distinguished between two types of agents who might be involved in change: individuals with adaptive preferences on the one hand, and organizations undertaking adaptive preference interventions on the other. Now, I want to address the question from another angle: rather than asking which agents might be involved in change, I want to ask how we should *understand* those agents – and in particular, how we should understand the agency of individuals with adaptive preferences being asked to make changes in their lives.

What I propose is that when we are asking whether we should want to try to change adaptive preferences, we must simultaneously ask two different questions about our agents of change:

(1) Will the person being asked to change their preference be made better off by changing that preference?
(2) Does that person have an other-regarding reason (or more strongly, an obligation), to change their preference?

The first question treats the person with adaptive preferences primarily as a patient, whose prudential interests must be taken seriously by those who are acting upon her. But the second question recognizes that the person with adaptive preferences is also a moral agent, who has moral responsibilities to others in addition to moral interests that others must consider.

The adaptive preferences literature has primarily focused on the first question, asking what kinds of changes are in the interests of the people being asked to change. That said, the literature has not entirely ignored the agency of those with adaptive preferences – and indeed, some very important work has been done to recognize that agency. For instance, Serene Khader has forcefully argued that defensible adaptive preference interventions must include those with adaptive preferences as equal partners in the diagnosis and response to adaptive preference, rather than as simply moral patients to be acted upon (Khader 2011). And many opponents of adaptive preference intervention have argued that those judged to have adaptive preferences are in fact competent moral agents making effective judgements about their own interests in difficult circumstances, rather than – in Narayan's wonderful turn of phrase – 'dupes of patriarchy' who have been hoodwinked into loving what harms them

(Narayan 2002; Baber 2007; Jaggar 2006). These focuses on agency in the adaptive preferences literature are both important and welcome.

What I want to note, however, is that the focus on agency that you find in the adaptive preferences literature tends to be of a particular kind: it is a focus on agential *capacities,* rather than a focus on agential *obligations*. While in the best cases people judged to have adaptive preferences are treated as agents whose judgements, capacities, and moral status must be taken seriously, it is still almost exclusively their own individual interests that are centred in discussions of whether and how we ought to want to change adaptive preferences.

But as we saw earlier, their interests are not the only ones at stake. Adaptive preferences are a problem because of the ways in which constricted circumstances can cause people to limit or contort what they pursue and desire. We should certainly care about the effects that oppressive circumstances have had on those whose preferences have already been limited or contorted – but in recognizing the role that circumstances play, we must also consider how current circumstances might shape those whose preferences are just beginning to form, and those whose preferences will form under them in the future. And given the recognition in the literature on victim's obligations to resist oppression of the ways in which individual choices are both shaped by *and also reinforce* oppressive systems, we must focus on the agency of those with adaptive preferences in a more capacious way.[7] We should certainly recognize that they are agents capable of rationally responding to their limited circumstances – but we must also recognize that they are agents who are answerable to others for the effects of their own behaviour. It is not enough to ask them whether their behaviour is compatible with their own flourishing, or how their circumstances might best be modified to enhance their own flourishing. We must also ask – and perhaps more importantly, persons with adaptive preferences must ask *themselves* – whether there are real costs that they must bear in order to reform the system that required them to compromise their own flourishing in the first place. A woman in a demeaning relationship might genuinely be made better off by staying for financial or other bundled reasons – yet there remains the very real possibility that she owes it to her daughter to end the marriage so that her daughter expects better for herself. The point here is not to establish whether this mother does or does not have an obligation to leave her partner. The point is rather that if we focus only on the question of whether

[7] In addition to the literature on victims' obligations to resist their own oppression, see Callan 2005 for the importance of considering the agency and obligations of those targeted by social justice interventions.

adaptive preferences benefit their holders, we miss out on the possibility of this obligation entirely. To see the tension here we must ask both whether adaptive preferences benefit their holders, *and* ask – indeed, ask in a capacious sense – whether we should nevertheless want to change them for other reasons.

In this paper I won't even begin to attempt to solve the problem of how to balance the prudential interests of those with adaptive preferences with the obligations that they might have to others. This is a very serious, and very difficult problem. The literature's current focus on the prudential interests of those being asked to change their adaptive preferences is supported by at least one very important consideration: those with adaptive preferences have already been harmed by an oppressive system, and that harm should absolutely play a mitigating role in the amount of cost that they can be asked to shoulder in the service of preventing future generations from being placed in the position to form harmful adaptive preferences.

But burdened moral agents remain moral agents. Mothers are the victims of patriarchy, but we remain answerable to our daughters for the choices that we make, for the ways in which we choose both to comply with and to undermine the oppressive structures that they will grow up within. So while I will not try to offer an answer here, if we are ever going to give a defensible answer to the question of whether bearers of adaptive preferences should want to change those preferences, we must at minimum begin by considering them in these dual roles: as at once both bearers of interests and bearers of obligations. And to recognize this, we must ask ourselves all four of our questions at once.

6. Conclusion

In this paper, I've clarified what is at stake when we ask ourselves questions about how and whether to respond to adaptive preferences. Much enormously important work has been done to consider the interests, the humanity, and the agency of those with adaptive preferences. But we need to do more to understand those with adaptive preferences as agents with obligations. I'm not going to try to say here how much we owe, or under what conditions we owe it. I will just insist that when we ask the

question of how to respond to adaptive preferences, we – as philosophers, as mothers, as other victims of oppression, as people working in governmental or non-governmental organizations – must understand adaptive preferences broadly in both mode and orientation, and think of those with adaptive preferences as both subjects with interests and as agents with obligations. And to get here, we must consider whether satisfying adaptive preferences is good for the people who have them, whether we should want to change them, who should be doing the changing, and what we should count as adaptive preferences in the first place. We must consider each of these questions, all at once, in light of each other. We must be willing to get messy.

Acknowledgements

For helpful feedback and lively discussion, I am grateful to audiences at the Minneapolis Workshop on Well-being and the workshop on Adaptive Preferences, Structural Injustice, and Moral Responsibility at the University of Zurich. I am also grateful to two anonymous reviewers for this journal.

ORCID

Rosa Terlazzo http://orcid.org/0000-0002-0764-3223

References

Ackerly, Brooke (2008) *Universal Human Rights in a World of Difference*. Cambridge University Press.

Baber, H. E. (2007) 'Adaptive preference' *Social Theory and Practice* **33**: 105–26. doi:10.5840/soctheorpract200733137.

Callan, Eamonn (2005) 'The Ethics of Assimilation' *Ethics* **115**: 471–500. doi:10.1086/428460.

Colburn, Ben (2011) 'Autonomy and Adaptive Preferences' *Utilitas* **23**: 52–71. doi:10.1017/s0953820810000440.

Cudd, Ann (2006) *Analyzing Oppression*. Oxford University Press.

Dorsey, Dale (2017) 'Adaptive Preferences Are a Red Herring' *Journal of the American Philosophical Association* **3**: 465–84. doi:10.1017/apa.2017.36.

Elster, Jon (1983) *Sour Grapes: Studies in the Subversion of Rationality*. Cambridge University Press.

Emerick, Barrett (2016) 'Love and Resistance: Moral Solidarity in the Face of Perceptual Failure' *Feminist Philosophy Quarterly* **2**: 1–21. doi:10.5206/fpq/2016.2.1.

Fuller, Lisa (2011) 'Knowing their own good: preferences and liberty in global ethics' in Thom Brooks, ed., *New Waves in Ethics*: 210–30. Palgrave Macmillan.

Gheaus, Anca (2023) 'Political Liberalism and the Dismantling of the Gendered Division of Labor' in David Sobel and Steven Wall, eds., *Oxford Studies in Political Philosophy Vol. 9*: 153–82. Oxford University Press.

Hay, Carol (2011) 'The Obligation to Resist Oppression' *Journal of Social Philosophy* **42**: 21–45. doi:10.1111/j.1467-9833.2010.01518.x.

Jaggar, Alison (2006) 'Reasoning about well-being: Nussbaum's method of justifying the capabilities approach' *Journal of Political Philosophy* **14**: 301–22. doi:10.1111/j.1467-9760.2006.00253.x.

Khader, Serene (2011) *Adaptive Preferences and Women's Empowerment*. Oxford University Press.

Khader, Serene (2012) 'Must Theorizing about Adaptive Preferences Deny Women's Agency?' *Journal of Applied Philosophy* **29**: 302–17. doi:10.1111/j.1468-5930.2012.00575.x.

Khader, Serene (2021) 'Self-Respect under Conditions of Oppression' in Richard Dean and Oliver Sensen, eds., *Respect: Philosophical Essays*: 232–51. Oxford University Press.

Nagar, Richa and Saraswati Raju (2003) 'Women, NGOs, and the Paradoxes of Empowerment and Disempowerment' *Antipode* **35:** 1–13. doi:10.1111/1467-8330.00298.

Narayan, Uma (2002/2018) 'Minds of their own: choices, autonomy, cultural practices, and other women' in Louise M. Antony and Charlotte E. Witt, eds., *A Mind of One's Own*: 418–32. Westview Press.

Nussbaum, Martha (2000) *Women and Human Development: The Capabilities Approach*. Cambridge University Press.

Schouten, Gina (2019) *Liberalism, Neutrality, and the Gendered Division of Labor*. Oxford University Press.

Sen, Amartya (1987) *On Ethics and Economics*. Blackwell.

Sen, Amartya (1999) *Development as Freedom*. Anchor Books.

Terlazzo, Rosa (2015a) 'Adaptive Preferences: Merging Political Accounts and Well-being Accounts' *Canadian Journal of Philosophy* **45**: 179–96. doi:10.1080/00455091.2015.1059252.

Terlazzo, Rosa (2015b) 'Conceptualizing Adaptive Preferences Respectfully' *Journal of Political Philosophy* **24**: 206–26. doi:10.1111/jopp.12062.

Terlazzo, Rosa (2018) 'Must Adaptive Preferences be Prudentially Bad for Us?' *Journal of the American Philosophical Association* **3**: 412–29. doi:10.1017/apa.2018.1.

Terlazzo, Rosa (2020) '(When) Do victims have obligations to resist oppression?' *Social Theory and Practice* **46**: 391–416. doi:10.5840/soctheorpract20204690.