



King's Research Portal

DOI:

[10.1093/mind/fzy020](https://doi.org/10.1093/mind/fzy020)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Mitchell, P. (2018). Adaptive Preferences, Adapted Preferences. *MIND*, 127(508), 1003–1025.
<https://doi.org/10.1093/mind/fzy020>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Adaptive Preferences, Adapted Preferences

POLLY MITCHELL

University College London

polly.mitchell@ucl.ac.uk

People who have not personally experienced diseases, disabilities and health conditions tend to judge them to be worse than they are reported to be by people who have experienced them (Ubel et al. 2005). This phenomenon, dubbed the *disability paradox*, presents a challenge for health policy. Given that assessments of the utility of health states are widely used as a measure health outcome, using patient, rather than public, assessments will radically change the distribution of healthcare resources.

This divergence between public and patient preferences is most plausibly explained as a result of hedonic adaptation (Loewenstein and Ubel 2008). Hedonic adaptation is a widespread phenomenon such that people tend to adapt fairly quickly to the state they are in, good or bad, and adjust their baseline utility accordingly. On the whole, after someone undergoes a major change in health or life circumstances, their self-reported quality of life drops or rises back to pre-event levels surprisingly quickly. Adaptation is not, however, widely recognised by people who haven't had the experience in question, nor indeed is it always recognised by people who *have* previously had the experience in question—hence the difference in public and patient assessment (Hausman 2015, pp.93–4).

One solution to the disability paradox is to show that patient utilities are inappropriate for use in public policy decision-making. This paper offers a critique of one such approach: the common assumption that adaptation leads to malformed or irrational preferences, sometimes called *adaptive preferences*.¹ I consider three philosophical accounts of adaptive preferences: Jon Elster's *procedural* account, Luc Bovens' *formal* account, Martha Nussbaum's *substantive* account. I argue that none of these accounts characterise adapted patient preferences as irrational. *Adapted* preferences should not be treated as synonymous with *adaptive* preferences. I suggest that much patient adaptation should be understood as a form of the ubiquitous human ability to respond to environmental change. Moreover, because they amount to testimonial denial, attributions of adaptive preferences should be made with caution. As such, we ought not to discount patient preferences.

1. The disability paradox

Health state utility values are widely used as a measure of health benefit, for calculating quality adjusted life years (QALYs), determining the cost-effectiveness of drugs and public health programmes, and determining the effectiveness of treatments in particular populations. They capture, in some sense, how good or bad it is to be in a particular health state. Typically, utility values are calculated by eliciting the preferences of a large sample of members of the public, who are roughly demographically representative of the overall population, but who do not necessarily have personal experience of the conditions in question. Public utility values therefore don't capture the quality of life associated with each health state, where this is a measure of what it's like to be in that state, but something like how much people want to avoid being in the state (Wolff et al. 2012, pp.460–1). Utility values can also be elicited by patients, who *do* have personal experience of the health conditions in question. In these cases, utility values also capture the extent to which patients

¹ See, for example, introduction to Raikkä and Varelius 2013.

want to be free from a particular health state, but from the perspective of someone who is currently experiencing it.²

Non-patients tend to overestimate the health losses and underestimate the health gains associated with many health states, in comparison to the reports of people with those health states. So, for example, in one study women without experience of breast cancer valued having and living with a mastectomy significantly lower than women who had experienced breast cancer, most of whom had undergone mastectomy (Ashby et al. 1994). In another, patients who had undergone colostomy as a result of surgery for rectal cancer assigned higher utilities to colostomy than did patients who had been treated for rectal cancer by radiotherapy without colostomy (Boyd et al. 1990). In such cases, patients report their condition to be less disabling, or have a smaller impact on their well-being, compared to the judgements of non-patients. This phenomenon, known as the *disability paradox* is widespread, observed across many different types of disease and disability (Ubel et al. 2005; Dolan and Kahneman 2008, pp.222–3).

Using utility values elicited by the public, rather than values elicited by patients who have the condition in question, will have a significant impact on the allocation of healthcare resources (Dolan and Kahneman 2008, pp.223–4; Loewenstein and Ubel 2008, p.1799). If we take the utility values elicited by patients to accurately represent their well-being, their possible health gains will be much smaller than if we use the utility values elicited by the public for the same conditions. Using higher utility values (typically patient values) means that health states will be taken to be less disabling, and to have less serious health detriments, and as such, curing them will result in fewer QALYs gained, or a smaller health benefit. Using lower values (typically public values) means that conditions represent a more significant health burden, and as such, curing them will result in more QALYs gained, or a larger health benefit. Public utility values will thus justify greater use of resources in their mitigation and prevention (Brock 2004, pp.204–5). Treatments which would have been deemed cost effective had the lower value been used may end up not being deemed cost effective using the higher value, resulting in their not being commissioned, or their deprioritisation (Ubel et al. 2003, p.600). This is particularly likely to be true in the case of physical disabilities, where adaptation is most evident, and mobility issues,³ where the discrepancy between public and patient values is large (Wolff et al. 2012, p.457).

This divergence constitutes a challenge for health policy decision-making. In the absence of further evidence to suggest that either patient or public assessments better represent how

² Measuring health state utility by eliciting preferences is a complex and controversial matter. There is some ambiguity about what exactly is being measured, and the presence of framing, focussing and anchoring effects calls into question whether such measures reliably capture people's preferences at all (see Hausman 2015, especially chapters 7 and 8). I will put these difficulties to one side in this paper. For the sake of my argument, I assume that the measurement of health state utility does capture, in some genuine sense, people's preferences.

³ In some cases, particularly mental health conditions, the effect may be reversed. Depressed people, for example, judge the quality of life associated with depression to be worse than the public perceive it to be (Pyne et al. 2009), and patients with dementia and their carers give systematically lower utility scores to dementia states than members of the general population (Rowen et al. 2015). In such cases, the use of public values is likely to result in *fewer* resources being used to prevent and treat these conditions, as they are seen as less disabling than they would be judged to be if patient values were used. This is not straightforwardly related to adaptation, and may be the result of stigma associated with some diseases, or failure to understand what they entail. Nonetheless, it presents similar problems for decision-making about healthcare resource distribution.

bad it is to be in a given health state, there is no clear answer as to whose values to use in cost-effectiveness analysis. On the one hand, non-patients generally have limited information about a given health state and what it's like to be in it, given that they have no lived experience of it. This may be a *prima facie* reason for preferring patient utility values as a measure of how good or bad it is to have a particular condition. However, public utility values may instead reflect, or at least partly reflect, people's fear about getting a condition, rather than their considered preferences about what it's like to be in that state (Dolan and Kahneman 2008). The lower utility values that the public ascribe to health states may thus register a genuine and widely shared preference that has relevance for health-policy decision-making. This would suggest that patient experience doesn't capture everything that is relevant for measuring utility in a public policy context (Wolff et al. 2012, p.460–1). Of course, some diseases and disabilities are highly stigmatised, and in these cases public preferences may reflect prejudicial attitudes. It may be inappropriate to make policy decisions on the basis of such preferences, if they are distorted by stigma.

Moreover, patients may misrepresent how bad it is to have a disease or condition, such that their valuation of their own health state is distorted or misguided in some sense. Some cases appear to support this claim. For example, one study showed that patients who had previously had a colostomy bag, now removed, valued the utility of living with a colostomy bag almost identically to people who had never had a colostomy bag. Moreover, both groups valued this state much lower than those currently experiencing living with a colostomy bag (Hausman 2015, pp.93–4). Given the reversion to the lower utility value, even having experienced living with a colostomy bag, cases like this may suggest that the patient's reported utility is mistaken. If patient preferences can be characterised as misguided or irrational, they will fail to capture utility accurately, and can for that reason be discounted from social choice functions. That being the case, the disability paradox would fade away.

2. Hedonic adaptation

The most plausible explanation of the divergence between public and patient assessment of disease and ill health is *hedonic adaptation*. Hedonic adaptation is a widespread feature of reported health-related quality of life. After a change in life circumstances, people tend to adapt fairly quickly to the state they are in, good or bad, and adjust their reported utility accordingly (Kahneman 2000). As such, reported utility or quality of life drops or rises back to pre-event levels surprisingly quickly. So, for example, accident survivors who have sustained paraplegia or quadriplegia report happiness levels only slightly lower than those reported by non-disabled control groups within a year of the event (Brickman et al. 1978). Patients who have lost limbs to cancer or suffered burn injuries report similar quality of life and psychiatric symptoms to control groups (Dolan and Kahneman 2008, p.218). Furthermore, for people who have experienced major negative changes in their health status in the past, the onset of new conditions is less likely to result in self-reports of worse health, indicating a resilience to new threats to quality of life (ibid.).

Adaptation also occurs in a non-health context: after receiving a pay rise or getting married, individuals' reported happiness levels significantly rise, but then return to their baseline within a short time period; after the death of a spouse, a widowed person's reported happiness levels return to their baseline after, on average, eight years (ibid.). At a population level, while real income in Japan increased fivefold between 1958 and 1987, self reported happiness did not change (Kahneman 2000, pp.686). While adaptation is common, it is not universal. For example, there is evidence that people do not adapt their utility levels in response to unemployment, and people exhibit *increased* sensitisation to annoying noise over

time (Frederick and Loewenstein 1999, p.311). In the health context, there are a number of cases where adaptation does not occur: for example, people suffering from long-term unexplained pain do not experience adaptation, with reported quality of life getting worse over time (Dolan and Kahneman 2008, p.218). Experiencing repeated episodes of pain may also make patients *less* resilient in the face of future stressful events. Adaptation does not appear to occur in degenerative illness, where a patient's health status is continually changing (Frederick and Loewenstein 1999, p.312).

Although hedonic adaptation is common, people largely fail to recognise its existence, expecting positive and negative experiences to last longer than they actually do, not only when judging how they will feel in the future but also when estimating other people's well-being following an episode of good fortune or ill health (Dolan and Kahneman 2008, p.221–223; Kahneman and Thaler 2006, pp.229–31). This accounts for the public judging health states to be worse than patients report them to be.

There are a number of proposed explanations for how and why adaptation occurs. Most focus around the idea of a *response shift*, that is, that people's internal standards, values, and conception of well-being change. Adaptation is described as a treadmill effect: people respond to changes in their health or life status by adjusting their evaluative standards, and as such the well-being they experience remains broadly stable around a baseline level (Brickman and Campbell 1971, p.287). This effect is demonstrated in an oft-cited study, which found that lottery winners rated their present happiness only slightly higher than controls, and people who had sustained paraplegia or quadriplegia in an accident rated their present happiness only slightly lower (Brickman et al. 1978, p.921). Both lottery winners and paraplegic people rated their past happiness as higher than members of the control group, which seems to indicate a nostalgia effect, where past experiences are remembered as better than they actually were experienced as being (ibid. p.924).

There are several ways in which people exhibiting adaptation are thought to change their standards of evaluation. For example, it has been suggested that paraplegic people compare their well-being and function to that of other people with paraplegia, rather than to the population in general, and that people who have experienced a major negative life event compare their current well-being with the very negative well-being experienced in the transition period immediately following the change, rather than comparing it with their pre-event experiences, in both cases artificially increasing their reported well-being (Ubel et al. 2003, p.604). Changing standards might, to some extent at least, be appropriate. Consider a 30-year-old and an 85-year-old, who both rate their well-being as 95 out of 100. It would be reasonable to expect that there would be significant differences in their physical and psychosocial function, and that their answers might reflect their expectations of best possible quality of life, given their age and context, rather than some external, contextless scale of well-being (ibid., p.602). However, such relativism may sometimes be inappropriate. For example, if someone with symptomatic heart failure uses other heart failure patients as a benchmark for their well-being, rather than normal age-appropriate adults, they may be in danger of failing to take into account the impact of their functionally limiting symptoms on their well-being. Other aspects of adaptation may include learning to cope with limited function by acquiring or developing skills, lowering expectations of feasible function and accomplishment, and changing opinions about what it takes to be happy, or to live a fulfilling life (Ubel et al 2003, p.604; Menzel et al. 2002, p.2152).

3. Formal and procedural adaptive preferences

If those preferences resulting from hedonic adaptation, or *adapted preferences*, can be shown to be mistaken, patient preferences could be justifiably discounted. One way of characterising the mistakenness of adapted preferences is by classifying them as *adaptive preferences*. Although in some sense all of our preferences develop and adapt in response to our environment, adaptive preferences are those preferences which have adapted in such a way as to make them irrational or otherwise unreliable in reflecting our actual attitudes (Begon 2015, p.243).⁴

A few examples will elucidate the nature of those preferences typically considered to be adaptive. Martha Nussbaum gives several compelling examples of women in oppressive personal and socio-economic contexts in India, for whom it seems obvious that the standards by which they measure their well-being or internal state are distorted, and their resulting preferences are problematically adaptive (2000a, pp.112–113). Vasanti, after years in an abusive marriage, thought her abuse to be a normal part of a woman's life, something to be expected once she left her family home to live with her husband. Jayamma, despite being paid less than men for more demanding factory work, accepted that this was how things were, and, knowing change was not possible, did not even waste energy lamenting her situation. And severely malnourished women in Andhra Pradesh, prior to the efforts of a government consciousness-raising program, didn't consider themselves to be malnourished, or their conditions to be unhealthy. Examples like these are widely reported. Women who experience violence at the hands of their partner are less likely to seek medical care, not only for any resulting injuries, but also preventative care such as cancer screening and blood pressure checks (Garcia-Moreno et al. 2012, p.5). People in abusive relationships sometimes grow to feel guilt and shame for their own actions, blame themselves for misstepping, and feel that they deserve punishment (O'Neill and Kerig 2000). People from oppressed ethnic minorities in racist societies experience losses in self-confidence, motivation and self-esteem, in response to their treatment as inferior (Nagel 1973, p.350).

Adaptive preferences are typically characterised by a number of features, including changes in self-perception; incorrect judgement about the appropriateness of the actions of others, or judgement against incorrect standards; and pursuit of self-destructive or self-deprecating ends. As patient adaptation is also characterised by changes in self-perception and changes in internal standards of judgement of oneself and others, it seems to be a *prima facie* candidate for a type of adaptive preference. Given that patient adaptation results in patients assigning high utilities to disease states, which can lead to a preference ranking of disease states over non-disease states at the personal level, and the downgrading of the health-related costs of particular diseases and disabilities in social choice functions, such adaptation can also plausibly be described as self-destructive.

There are three principal accounts of adaptive preferences that ostensibly give credence to the view that patient preferences should not be used in utility functions: (1) Jon Elster's

⁴ Sometimes adaptation and adaptive preferences are treated as synonymous. For example, many of the contributors to Juha Räikkä's and Jukka Varelius' comprehensive edited volume on the subject (2013) appear to use the two concepts interchangeably. I start with the assumption, shared by Jessica Begon (2015), that not all adaptation is adaptive. If the two are *not* synonymous—as I argue in this paper—adaptive preferences will form only a subset of adapted preferences, that is, those adapted preferences which are irrational or misguided. The terminological resemblance of 'adapted preferences' and 'adaptive preferences' is, perhaps, unfortunate. However, it serves the purpose of highlighting the close relationship between the two, even if, as I will argue, they turn out to be non-identical.

procedural account, which takes the problem with adaptive preferences to be that they come about in the wrong way; (2) Luc Bovens' *formal* account, according to which the problem is the conflict between adaptive preferences and other preferences; and (3) Martha Nussbaum's *substantive* account, which contends that adaptive preferences, which are typically formed in deprived contexts, are in conflict with goods which have intrinsic value. In this and the following section, I consider each in turn. I argue that in fact none of these accounts of adaptive preferences show adapted patient preferences to be adaptive preferences.

Jon Elster (1983) argues that adaptive preference change is non-autonomous and so irrational. Elster distinguishes adaptive preference formation from deliberate character planning, both of which occur in the context of reduced possibilities: the former is a causal, non-conscious process motivated by a drive to reduce the frustration that results from unsatisfied desires; the latter is the deliberate, intentional shaping of desires. For Elster, the prototypical example of adaptive preferences comes from La Fontaine's fable of the fox and the grapes. Upon realising that he cannot reach the grapes that he so desires, and rather than admit defeat, the fox resolves that they look sour and that he didn't want them anyway. In contrast, a character planning fox, realising that he cannot reach the grapes and finding only sour cherries within his reach—usually not sweet enough for his taste—decides to try to learn to like the cherries. After a while, he finds that he actually prefers the tartness of the cherries to grapes: he has brought about a preference change in himself. These are both examples of cognitive dissonance reduction: rather than live with an unsatisfied desire, the fox changes his desires. The difference, according to Elster, is that whereas the adaptive fox undergoes a change in preference motivated by a non-autonomous subconscious drive, the character planning fox undergoes a genuine change in preference, motivated by an autonomous second-order preference. The first fox's alleged preference change seems insincere and resentful, and would undoubtedly be reversed if he found a way to reach the grapes, whereas the second fox's preference change is genuine and stable. Elster also describes adaptive preferences as typically downgrading inaccessible options, whereas deliberate character planning typically upgrades accessible ones.

Luc Bovens (1992) criticises Elster's account of the mechanism of adaptive preferences, though he broadly agrees that sour-grapes-type situations tend to lead to irrational preferences, and character planning tends to result in rational preferences. He argues that Elster's distinction between intentional, conscious meta-preferences about what your preferences should be and unintentional, unconscious, mechanistic drives does not in fact track the distinction between rational and irrational preferences. There are many parts of our identity which aren't consciously decided—sexual preferences perhaps, or music taste—but which are not thereby irrational. Donald Bruckner similarly suggests that non-conscious adaptation to changed circumstances can be conducive to subjective well-being, and that to adopt such preferences is more rational than not (Bruckner 2009, p.315). Instead, for Bovens, what makes a changed preference irrational is a conflict with other preferences of the same type that the agent continues to hold. So, when the fox states his preference not to eat the bunch of grapes because they look sour, but retains a general preference to eat grapes that look similar, his token preference not to eat the grapes contradicts his type preference to generally eat similar grapes. The character planning fox, on the other hand, has rational preferences: his preferences towards the types 'sweet fruits' and 'sour fruits' have changed such that his token preference for sour cherries over grapes is not in tension with his overall preference set. Other accounts of adaptive preferences give a similarly formal explanation of their irrationality: Bruckner (2009), for example, argues that irrational preferences are those that are not reflectively endorsed by the agent; John Christman (2014) suggests that

autonomous preferences are those that are reflexively endorsed by an agent, and form part of a self-affirming practical identity.

If the preference changes that patients undergo are generated by the kind of non-intentional drive that Elster describes in the case of the adaptive fox, or if they are in conflict with general preference types, as Bovens outlines, then describing the lower utility values reported by patients adaptive preferences may be justified. For Elster, patient preferences which develop unconsciously, non-intentionally, and are motivated out of a drive to resolve the frustration of not being able to satisfy previously held preferences, would be classed as irrational, and so as adaptive preferences. For Bovens, patient preferences which conflict with their other preferences—for example, a patient with chronic kidney disease who reports minimal change in quality of life despite having to attend 12 hours of haemodialysis per week, but who also reports a strong preference for time spent out of hospital over time spent in hospital—would be similarly irrational.

Though neither Elster nor Bovens discuss changes in health-related preferences directly, there is reason to think that neither the procedural nor formal description of irrational preferences in their accounts fully explain adapted patient preferences about health states. A detailed characterisation of the mechanisms of adaptation (Menzel et al. 2002) identifies a variety of changes and mechanisms which seem to point towards far more complexity in patient adaptation than either Elster or Bovens' accounts allow for. Elements of adaptation may include cognitive denial of functional health state, suppressed recognition of full health, skill enhancement, activity adjustment, substantive goal adjustment, altered conception of health, lowered expectations and heightened stoicism (*ibid.*, pp.5151–2). Differences in patients' reported wellbeing immediately after contracting a disease or health condition and after they have lived with the condition for a few months or years, are understood to be down to a combination of these factors.

Some of these aspects of adaptation are roughly in line with Elster's suggested mechanism for adaptive preferences. For example, when patients exhibit cognitive denial of functional health—wherein they find it hard to admit how bad their functioning really is—or suppressed recognition of what full health would be like and what their current health state prevents them from doing, their adapted preferences look much like Elster's adaptive preferences. Another feature of adaptation that is characteristic of Elster's sour grapes account is when patients lower their expectations about their own well-being and function, which may represent a regrettable acceptance of a diminished life. Such features of adaptation may also be taken to lead to irrational preferences on Bovens' account, for example, if cognitive denial, modified benchmarks and lowered expectations are confined to preferences relating only to patients' health, or relate only to the particular condition from which they suffer and the ways in which it limits their behaviour and options. In such cases there may be contradictions between specific disease- and behaviour-related preferences and general preferences of the same type.

However other, more positive, elements of patient adaption are less obvious candidates for adaptive preferences. For example consider patients who adjust their activities and substantive goals, taking up new interests and activities which are allowed or enabled by their impairment, or people who enhance their skills, either with respect to old or new goals and activities, as a result of their ill health or disability. Some of these changes in commitments might be motivated by a foreclosure of other options, for example, someone who gains a newfound interest in creative writing because they are not able to engage in their previous love of cycling, having lost the use of their lower limbs. This might constitute a substantial change in goals. However other activities might be enabled by impairment, for example

participation in disability sports, which are only open to those with physical or developmental impairments, or new friendships and membership via patient support networks. Sometimes such adaptation might constitute a change in activity which realises a consistent goal via a different means: success in sport, for example. This would not necessarily imply a change of preferences at all: rather, ill health or disability may necessitate changes in the means to satisfaction of a patient's preferences. Here adaptation appears to be a conscious and intentional decision, motivated partly by the foreclosure of some options, partly by the new accessibility of others: perhaps more like character planning than sour grapes. Furthermore, adaptation may not involve conflicting preferences, if it amounts to a comprehensive change of preferences and goals.

People with long term conditions and disabilities may also come to adopt a view that health is not identical with high levels of physical functioning, and come to hold a more holistic account of health—understood as the ability to live a full, purposive and valuable life, rather than simply a matter of biological function—or they may realise that they can achieve happiness without full health. This does not necessarily represent lower expectations, but a change in priorities. After the onset of an illness, a patient's internal standard for evaluating their own health may change: what previously was taken to constitute a 0.5-type health state, might, on reflection, and with experience, be reevaluated as a 0.75-type state instead (Ubel et al. 2003, p.602). This might be best understood as neither sour-grapes-type adaptive preferences nor as deliberate character planning, but something more like what Elster calls 'learning and experience' (1983, pp.113–5). Sometimes the status quo is preferred for no reason other than its being the status quo, or because other options have not been experienced, or because change itself is undesirable, and it takes experience of other ways of living to agitate stagnated preferences. Experience of new options and different lifestyles can be an opportunity for people to come to know their value, resulting in changed preferences. This is different from character planning, in that an agent does not necessarily set out to change her preferences and may resent the change at first, and different from adaptive preferences, insofar as the preference change is informed and represents increased knowledge of the option-set. Preference change through learning is not, for Elster, irrational, because it generates informed preferences.

Patient adaptation looks, at best, to be a lot more complex than the narrative Elster and Bovens tell us about adaptive preferences, and to comprise, at least in part, a rather different phenomenon. If reported increase in wellbeing is partly a matter of shifting benchmarks, partly a matter of changed, rather than diminished, interests and goals, and partly a result of a changed outlook on what matters in life, it's not clear whether the resulting preferences should be considered adaptive or merely different. Bovens discusses the existence of a large grey area between clear instances of sour-grapes and character planning (Bovens 1992, pp.76–7). He takes there to be reasons for including cases that appear somewhere in the middle of the continuum to be classed as either sour-grapes or character-planning, without much indication of how to make this call, but ultimately thinks that many cases will remain undecided. If patient adaptation sits in this grey area, as seems plausible, neither Elster nor Bovens shed much light on the question of whether they are problematically adaptive or not. Notably the formal accounts of adaptive preferences which take reflective endorsement to defeat claims of irrationality (Bruckner 2009; Christman 2014) entail that much adaptation, including patient adaptation, is rational and should be taken to reflect genuine preferences.

4. Substantive adaptive preferences

Substantive critiques of adaptive preferences suggest that the problem lies not with the process of preference formation, nor the consistency of preference sets, but with the *content* of certain preferences. Martha Nussbaum's account, for example, is explicitly normative, identifying preferences as irrational when they favour those outcomes which *ought not* to be preferred (2001a, p.149). The goods that it is rational to prefer are specified by her list of ten central capabilities; preferring not to have an item on the list precludes a person's preferences from inclusion in a social choice function. For Nussbaum, the items on the list constitute those capabilities which are necessary for human functioning that is 'worthy of the dignity of the human being', and their pursuit is thus normatively justified (*ibid.*, p.148). On this account, adaptive preferences are irrational not because they judge well-being against a changed standard *per se*, but because they express a preference for something which ought not to be preferred. Typically such adaptation occurs in response to conditions of extreme deprivation or oppression. Nussbaum thinks that any account of adaptive preferences relies on a substantive theory of the good, and the capabilities approach offers a candid and reasoned account thereof.

If patient utility values represent preferences for suboptimal outcomes, they too will count as adaptive preferences on Nussbaum's account, and thus be ruled out from use in public policy decisions. Patients value, in some cases quite highly, states of being which include significantly reduced physical and psychosocial function, and as such adapted patient preferences violate, most plausibly, the capability for 'bodily health' (Nussbaum 2001b, p.87). Serene Khader's perfectionist account of adaptive preferences appears to have similar implications. She argues that adaptive preferences fulfil three conditions: (1) they are inconsistent with a person's basic flourishing; (2) they are formed under conditions non-conducive to her basic flourishing; and (3) they would not have formed under conditions which are conducive to basic flourishing (Khader 2011, p.51). As it is very likely that a patient's preferences would not have adapted had she not become ill, if her ill health is inconsistent with her basic flourishing, then her adapted preferences will be adaptive in this sense. While my discussion in this section engages with Nussbaum's account, my arguments equally apply to other substantive accounts of adaptive preferences, such as Khader's.

In order to show that patient preferences are adaptive, Nussbaum would need to show that fulfilment of the capability for bodily health precludes such preferences. However this is exactly the question under scrutiny: whether patient-reported utility is a good measure of health-related benefit. If health is understood entirely in terms of functional status and clinical evaluation, then there may be good reason to reject patient claims that their quality of life is higher than others anticipate. However there is evidence that an account of health should also incorporate subjective measures of value, namely, patient reports that 'objectively' bad health isn't as bad as supposed, or at least doesn't have as great an impact on quality of life as supposed, after all. To reject patient testimony on the grounds that it contradicts objective measures of health begs the question, if the existence of the contradictory testimony serves to question the claim that health should be measured objectively.

Nussbaum *does* want to afford some space for subjective preferences or desires in her account of well-being. First, she accepts that some individuals will reject the goods afforded by her theory, and is relatively untroubled by this, as long as this rejection takes place at the level of functioning rather than the level of capability (2001a, p.154). That is, everyone is free to choose not to pursue these goods, despite having the opportunity to do so, but not to deny them to others or to remove their right to them. Nussbaum also argues that 'informed desire' should play a role in generating the values in her substantive list of capabilities, in

order to ensure the set of values is one which people with diverse lifestyles and preferences will live by (ibid., p.152). Her normative list is not, then, supposed to be the result of armchair theorising, but to reflect a diversity of subjective preferences. Nussbaum claims that attending to what people desire is part of what it is to respect a person—to reject their subjective experience out of hand is to fail to treat them with dignity (ibid., p.154). Given this, Nussbaum’s account cannot straightforwardly rule out the testimony of patients with adapted preferences. Patient-reported quality of life values, for example, may provide evidence that helps to define the limits of the concept of the capability for bodily health. Specifically, they suggest that a purely clinical and functional account of health fails to account for many patient’s experience of ill health, and that the inclusion of subjective metrics into a definition of health may be necessary for capturing it adequately. In order to deny this, some other reason—that is, *not* a restatement of the commitment to an objective account of health—needs to be given to undermine the validity of patient testimony.

In her discussion of adaptive preferences in the context of disability, Elizabeth Barnes argues that the bar for diagnosing adaptive preferences should be set very high, and that unsubstantiated diagnosis of adaptive preferences is a form of testimonial injustice (2009a; 2009b; 2016). Although the situation of disabled people is not identical to that of people in states of ill health, her argument has relevance here.⁵ Barnes argues that to discount the positive testimony of disabled people is to claim that the non-disabled majority is in a better position to determine the well-being of disabled people than they are themselves (2016, p.133). Given that it is usually assumed that for any two people, all things being equal, *x*’s testimony is preferable to *y*’s testimony as evidence of *x*’s well-being, we need a good reason to think that things are not equal, and that *x*’s testimony is misrepresentative (2009a, p.10). This cannot be the claim that disability is a suboptimal state, as this is exactly the matter up for debate. This looks very similar to the problem of patient adaptation. Unlike the case of disability, patients do not usually claim that their state of ill health is as valuable to them as full health; instead they make the weaker claim that their state of ill health is not as bad as other people think it to be. However, at issue in both cases is the testimony of some individuals about their subjective state, which contradicts the judgements of others.

One reason for which we might deem it appropriate to discount the preferences of patients and disabled people is because, even when they have had past experience of full health and functional range, their range of options is limited, and in some cases limited in such a way that cannot be changed. It is psychologically advantageous for people with limited options to adapt their preferences to avoid cognitive dissonance, but these coping mechanisms mask lower levels of well-being (Barnes 2009b, p.344). Barnes points out that the fact that a state constrains options, or involves functional limitations, even permanently, doesn’t itself make preferences formed within that state adaptive (2016, p.59). Being born with male sex organs entails not being able to bear and give birth to children. However we

⁵ The distinction between disability and ill health, particularly chronic and irreversible ill health, is not a clean one, and at least partly a matter of convention. I maintain the distinction for the sake of this discussion, mainly due to differences in the way in which well-being is measured in the set of people Barnes and I are respectively concerned with. Barnes focuses on disabled people’s testimony that their lives are as valuable as those of non-disabled people, whereas my main concern here is patients reporting their quality of life through some kind of preference elicitation tool, which often results in higher-than-expected utility values. There may well be significant overlap between these two groups, but claiming that your life is as valuable as someone else’s need not entail the claim that you experience the same level of utility or the same quality of life as them.

don't think of people with male sex organs as having problematically constrained options due to this inability, or consider their well-being to be lower because of it.

Indeed, *everyone's* goals, pursuits and expectations are developed within a context of their functional limitations (Barnes 2016, p.132). I don't compare my sporting ability with an olympic athlete, nor do I judge my strength-to-weight ratio against that of an ant. Furthermore, my goals and life plans are constructed roughly with the expectation that I will live an averagely long and disability-free life, not in the expectation that I will live to 150 years, or die in an accident tomorrow. My goals, pursuits, and expectations of my ability are relativised to my own capacity and context, which includes things like my species membership, my physical function and capacity, but also the behaviour of others around me, the set of people to whom I take it to be relevant to compare my welfare, and my understanding of my own capabilities. If my capacity or context changes in some relevant way, it is seemingly appropriate for my behaviour and preferences to change as well, or at least to be liable to change. If there isn't good reason to think that it is only the internal standards of patients that adapt according their context, then there doesn't seem to be good reason to take patient adaptation to be peculiarly problematic. Characterising patients as misguided about their reported utility on the basis of their functional limitation might entail very restrictive claims about people's ability to report their own subjective states in general, given the universality of functional limitations. This is consonant with the literature on hedonic adaptation. Adaptation is typically taken to be a ubiquitous phenomenon, something which all people experience, rather than something reserved for exceptional instances of loss or gain (Brickman and Campbell 1971, p.289). While disease and disability may present particular questions for adaptation—for example if they generate sudden downwards shifts in affect, rather than gradual change—perhaps they aren't categorically different from other instances of adaptation (*ibid*, p.292).

However, the limited function of disabled people and patients may differ from 'ordinary' limited function. One such difference could be that disabled people and patients don't know what they're missing out on, and so mis-value their experiences by failing to make a salient comparison with non-disabled or non-diseased states (Barnes 2016, p.120). Of course this often not true in the case of disability—most disabled people have non-congenital disabilities—and is even less likely to be true in the case of ill health, the vast majority of which is acquired. It's furthermore unclear that non-disabled and healthy people are in a better position than disabled people and patients in this respect, unless they have had past experience of the disability or condition in question themselves. Moreover, however, typically we don't require people to know what it's like to be in a different state to give positive testimony (or any testimony at all) about what it's like to be in that state. For example, it's appropriate for a woman to give testimony about the value of being a woman and about her quality of life as a woman, despite not having identified as a man or as genderqueer at any point in her life. Even when people do not have experience of any other states of health or disability, there may not be sufficient reason for discounting their testimony about their own well-being.

Another possible distinguishing characteristic of the functional limitations of disabled people and patients is that, if they have a *mere lack* of functions and abilities rather than a different set of functions, such that the options they are left with are simply a subset of the functions of non-disabled, non-diseased people, there may be reason to think that their preferences are malformed. The claim here is that having fewer options is necessarily worse than having more options to choose from: functional absences shouldn't be confused with functional differences. This is pretty implausible in the case of disability: while some

disabilities do involve a lack (blindness, deafness, paralysis), compelling testimony from disability activists tells us that disability also gives rise to new abilities, liberation from cultural norms, and unique positive experiences. Blind people, for example, cannot judge people on the basis of how they look, and so sidestep some of the prejudicial attitudes that sighted people cannot help but adopt; congenitally deaf people have the experience of having signing as a first language (Barnes 2016, p.95). In the case of illness, however, this objection is somewhat more plausible. Unlike in the case of disability, where many disabled people report not only valuing their disabled existence, but say that they would rather be disabled than non-disabled (ibid., p.119), it's much less easy to identify the positive aspects of ill health, and most patient-reported well-being *is* lower than the reported well-being of non-patients.

However, clearly not all instances of reduced function—even when they can be understood as a mere lack—give us reason to discount someone's testimony about their well-being. The fact that I cannot run as fast as an olympic sprinter, nor play the violin as well as the leader of the London Symphony Orchestra—even if I am functionally equivalent to them in other respects—and the fact that these are not plausible career options for me, does not discount my claim that I value being a philosopher, nor my reports of my own well-being. If some types or degrees of functional limitation *are* reason to discount testimony about well-being, while others are not, some specification of why and when this occurs is needed. And of course, this cannot make reference to the inherent sub-optimality of ill health *per se*, as this is precisely the matter in question. To claim that the lack of some health-related functions inherently compromises or discounts the testimony of patients begs the question as to whether patient preferences should be discounted.

Given that patient reports of quality of life themselves provide evidence to suggest that ill health is not the deeply suboptimal outcome that it is sometimes supposed to be, the substantive model of adaptive preferences can't explain the irrationality of the adapted preferences of patients without begging the question. Nussbaum's substantive approach, like Elster's and Bovens' accounts, fails to show that patient adaptation is adaptive in a problematic way. Classifying all patient preferences as irrational would require a characterisation of ill health as suboptimal, which assumes the deficiency of the very testimony under consideration.

5. Concluding remarks

My discussion has not attempted to critique or vindicate any of the three accounts of adaptive preferences discussed. However, I have suggested that, on any of these accounts, it is at best indeterminate whether patient adaptation amounts to adaptive preference formation. A detailed evaluation of patient adaptation seems to point to a complex picture, including some negative and some positive changes, but it's unclear that these are any different from the ways in which any person's preferences respond to changes in their physical and psychological function, with age, for example, or the ways in which all humans—all biological organisms—respond to changes in their environment.

I don't intend to rule out the possibility of adaptive preferences being determinable in other contexts, though I think Barnes is right to say that the bar for attributing adaptive preferences should be set very high. To say that someone's preferences are incorrect or misguided is to say that other people can determine their preferences better than they can themselves. Unjustified testimonial denial is a serious injustice, so great care should be taken to ensure that the preferences in question really are defective. However, prototypical instances of adaptive preference formation typically have characteristics which distinguish

them from the kinds of hedonic adaptation that occurs in patients and the disabled. Adaptive preferences in contexts of domestic abuse and forced incarceration, for example, are linked to gas-lighting and psychological manipulation, which engender perceived inferiority. Other kinds of adaptive preferences, those arising in deeply sexist and racist contexts, for example, involve structural inequality and systematically differential treatment according to oppressive norms, leading to similar psychological effects.

Certainly there is significant stigmatisation of many physically and cognitively impaired people, which is liable to engender feelings of shame and low status, and might point to some instances of adaptive preferences in disabled people and patients. However, most illnesses and patients are not stigmatised in this way, and it's implausible to suppose that people develop such a sense of inferiority simply due to the limitations brought on by having a disease or disability. Where stigmatisation of patients and people with disabilities does occur to the extent that patients are led to question their own evaluative standpoint and their capacities as a judge of their own actions and those of others, this may well constitute a good reason for questioning the representativeness of their preferences. Of course, it would be surprising to think that this should lead to the rejection of the patient perspective in favour of a public perspective, the latter presumably being the source of the stigmatisation in question.

Adapted preferences are not identical to adaptive preferences. As such, adapted patient preferences cannot easily be characterised as misguided or irrational, nor can they be excluded from social choice functions, and the disability paradox remains unsolved. This means that patient preferences remain salient in decision-making about healthcare resource distribution, though it doesn't entail that patient preferences are necessarily the right set of preferences to use in all assessments of cost-effectiveness and health benefit. Sometimes public preferences, which represent apprehension about entering into a health state, rather than any experience of being in that state, may be appropriate, especially, perhaps, in decision-making about preventative healthcare and public health. However, adaptation does not alone undermine the authenticity of preferences. Furthermore, adaptation should not be regarded as an exceptional phenomenon: all preferences are contextually anchored to some extent, and as such responsive to environment and circumstantial change.⁶

References

- Ashby, J., M. O'Hanlon and M.J. Buxton 1994, 'The time trade-off technique: How do the valuations of breast cancer patients compare to those of other groups?', *Quality of Life Research* 3: 257–265.
- Barnes, Elizabeth 2009a, 'Disability and Adaptive Preference', *Philosophical Perspectives* 23(1): 1–22.
- 2009b, 'Disability, Minority, and Difference', *Journal of Applied Philosophy* 26(4): 337–355.
- 2016, *The Minority Body: A Theory of Disability*, Oxford: Oxford University Press.
- Begon, Jessica 2015, 'What are Adaptive Preferences? Exclusion and Disability in the Capability Approach', *Journal of Applied Philosophy* 32(3): 241–257.
- Bovens, Luc 1992, 'Sour Grapes and Character Planning', *The Journal of Philosophy* 89(2): 57–78.

⁶ Thanks to James Wilson for comments on earlier versions of this paper.

- Boyd, Norman F., Heather J. Sutherland, Karen Z. Heasman, David L. Tritchler and Bernard J. Cummings 1990, 'Whose Utilities for Decision Analysis?', *Medical Decision Making* 10(1): 58–67.
- Brickman, Philip and Donald Campbell 1971, 'Hedonic relativism and planning the good science', in Mortimer H. Appley (ed.), *Adaptation-Level Theory: A Symposium*, New York: New York Academic Press.
- Brickman, Philip, Dan Coates and Ronnie Janoff-Bulman 1978, 'Lottery Winners and Accident Victims: Is Happiness Relative?', *Journal of Personality and Social Psychology* 36(8): 917–927.
- Brock, Dan W. 2006, 'Ethical Issues in the Use of Cost Effectiveness Analysis for the Prioritisation of Health Care Resources', in Sudhir Anand, Fabienne Peter and Amartya Sen (eds.), *Public Health, Ethics, and Equity*, Oxford: Oxford University Press.
- Bruckner, Donald W. 2009, 'In defense of adaptive preferences', *Philosophical Studies* 142: 307–324.
- Christman, John 2014, 'Coping or Oppression: Autonomy and Adaptation to Circumstance' in Andrea Veltman and Mark Piper (eds.), *Autonomy, Oppression, and Gender*, Oxford: Oxford University Press.
- Dolan, Paul and Daniel Kahneman 2008, 'Interpretations of Utility and their Implications for the Valuation of Health', *The Economic Journal* 118: 215–234.
- Elster, Jon 1983, *Sour Grapes: Studies in the subversion of rationality*, Cambridge: Cambridge University Press.
- Frederick, Shane and George Loewenstein 1999, 'Hedonic Adaptation' in Daniel Kahneman, Ed Diener and Norbert Schwarz (eds.), *Well-Being: Foundations of Hedonic Psychology*, New York: Russell Sage Foundation.
- Garcia-Moreno, Claudia, Alessandra Guedes and Wendy Knerr 2012, 'Understanding and addressing violence against women: Health consequences', *World Health Organisation*, <http://apps.who.int/iris/bitstream/10665/77431/1/WHO_RHR_12.43_eng.pdf> (accessed 24 April).
- Hausman, Daniel M. 2015, *Valuing Health: Well-Being, Freedom, and Suffering*, Oxford: Oxford University Press.
- Kahneman, Daniel 2000, 'Experienced Utility and Objective Happiness: A Moment-Based Approach' in Daniel Kahneman and Amos Tversky (eds.), *Choices, Values, and Frames*, New York: Cambridge University Press.
- Kahneman, Daniel and Richard H. Thaler 2006, 'Anomalies: Utility Maximisation and Experienced Utility', *Journal of Economic Perspectives* 20(1): 221–234.
- Khader, Serene J. 2011, *Adaptive Preferences and Women's Empowerment*, Oxford and New York: Oxford University Press.
- Loewenstein, George and Peter A. Ubel 2008, 'Hedonic adaptation and the role of decision and experience utility in public policy', *Journal of Public Economics* 92: 1795–1810.
- Menzel, Paul, Paul Dolan, Jeff Richardson and Jan Abel Olsen 2002, 'The role of adaptation to disability and disease in health state valuation: a preliminary normative analysis', *Social Science & Medicine* 55: 2149–2158.
- Nagel, Thomas 1973, 'Equal Treatment and Compensatory Discrimination', *Philosophy & Public Affairs* 2(4): 348–363.
- Nussbaum, Martha C. 2001a, *Women and Human Development: The Capabilities Approach*, New York: Cambridge University Press.
- 2001b, 'Adaptive Preferences and Women's Options', *Economics and Philosophy* 17(1): 67–88.

- O'Neill, Melanie L. and Patricia K. Kerig 2000, 'Attributions of Self-Blame and Perceived Control as Moderators of Adjustment in Battered Women', *Journal of Interpersonal Violence* 15(10): 1036–1049.
- Pyne, Jeffrey M., John C. Fortney, Shanti Tripathi, David Feeny, Peter Ubel, and John Brazier 2009, 'How Bad Is Depression? Preference Score Estimates from Depressed Patients and the General Population', *Health Services Research* 44(4): 1406–1423.
- Räikkä, Juha and Jukka Varelius (eds.) 2013, *Adaptation and Autonomy: Adaptive Preferences in Enhancing and Ending Life*, New York: Springer.
- Rowen, Donna, Brendan Mulhern, Sube Banerjee, Rhian Tait, Caroline Watchurst, Sarah C. Smith, Tracey A. Young, Martin Knapp and John E. Brazier 2015, 'Comparison of General Population, Patient, and Carer Utility Values for Dementia Health States', *Medical Decision Making* 35(1): 68–80.
- Ubel, Peter A., George Loewenstein and Christopher Jepson 2003, 'Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public', *Quality of Life Research* 12: 599–607.
- Ubel, Peter A., George Loewenstein, Norman Schwartz and Dylan Smith 2005, 'Misimagining the unimaginable: The disability paradox and health care decision making', *Health Psychology* 24(4 Suppl.): S57–62.
- Wolff, Jonathan, Sarah Edwards, Sarah Richmond, Shepley Orr and Geraint Rees 2012, 'Evaluating Interventions in Health: A Reconciliatory Approach', *Bioethics* 26(9): 455–463.