First among equals? Adaptive preferences and the limits of autonomy in medical ethics

Susan Pennings,[1] Xavier Symons[2]

[1]School of Philosophy, Australian National University, Canberra, Australian Capital Territory, Australia
[2]Plunkett Centre for Ethics, Australian Catholic University and St Vincent's Health Australia, Sydney, New South Wales, Australia

Correspondence to
Dr Susan Pennings, School of Philosophy, Australian National University, Canberra, ACT 0200, Australia;
susan.pennings@gmail.com

Received 19 October 2021
Accepted 31 January 2022
Published Online First
17 February 2022

To cite: Pennings S, Symons X. *J Med Ethics* 2024;**50**:212–218.

ABSTRACT

Respect for patient autonomy is a central principle of medical ethics. However, there are important unresolved questions about the characteristics of an autonomous decision, and whether some autonomous preferences should be subject to more scrutiny than others.

In this paper, we consider whether *inappropriately adaptive preferences*—preferences that are based on and that may perpetuate social injustice—should be categorised as autonomous in a way that gives them normative authority. Some philosophers have argued that inappropriately adaptive preferences do not have normative authority, because they are only a reflection of a person's social context and not of their true self. Under this view, medical professionals who refuse to carry out actions which are based on inappropriately adaptive preferences are not in fact violating their patient's autonomy. However, we argue that it is very difficult to articulate a systematic and principled distinction between normal autonomous preferences and inappropriately adaptive preferences, especially if this distinction needs to be useful for clinicians in real-life situations. This makes it difficult to argue that inappropriately adaptive preferences are straightforwardly non-autonomous.

Given this problem, we argue that there are significant theoretical issues with contemporary understandings of autonomy in bioethics. We discuss what this might mean for the practice of medicine and for medical ethics education.

INTRODUCTION

Respect for autonomy is a basic principle of medical ethics. In some instances, autonomy has become *the* dominant principle which people use to understand the obligations of doctors to their patients and medical researchers to research participants.[1 2] There are, however, significant questions surrounding what ought to be considered an autonomous decision and whether some autonomous preferences should be subject to more scrutiny than others. In this paper, we discuss the notion of *adaptive preferences* with a particular focus on preferences shaped by unjust and discriminatory social norms. We consider whether *inappropriately adaptive preferences*—preferences that are based on and that may perpetuate social injustice—should be categorised as autonomous. We argue that there is no straightforward criterion according to which inappropriately adaptive preferences can be deemed to be non-autonomous. In light of this, we argue that there are significant and unresolved theoretical issues with contemporary understandings of autonomy in bioethics. We also provide suggestions of how to respond to the challenge of inappropriately adaptive preferences in practice.

AUTONOMY IN MEDICAL ETHICS

Respecting a patient's autonomy has long been considered to be an essential part of ethical conduct in medicine and biomedical research. Beauchamp and Childress's influential principlist framework for bioethics places autonomy at the centre of medical ethics, along with the principles of beneficence, non-maleficence and justice. Beauchamp argues that the 'moral value of respect for autonomy precedes and is not the product of a philosophical theory, and no theory is acceptable if it conflicts with this value'.[3]

The principle of respect for autonomy is closely linked to the practice of informed consent in medicine. Informed consent procedures are thought to be essential because they safeguard patient autonomy. People should be able to choose the treatment options they wish to pursue, and choose whether or not they want to participate in research. Clinicians and researchers should refrain from placing any undue influence on patients or research participants as they make these decisions. This kind of respect is sometimes described as a 'constraint' on biomedical practice. In many cases, a patient's preferences about their care—such as whether they want to participate in a clinical trial, whether they want palliative care at the end of life, whether or not to take a medication or have a surgery—may conflict with what a doctor recommends. Patient preferences must, however, be respected. To do otherwise would be to constrain the autonomous choices of patients.[i]

In contemporary discussions, other major principles of biomedical ethics are often subsumed under the category of autonomy. For example, M. Therese Lysaught argues that beneficence has come to be defined with reference to the autonomous choices of an agent. According to Lysaught, one contemporary definition of beneficence would be 'the principle wherein we are obligated 'to do the other their good,' or to do good to the other as he/she defines it'.[4] This is significantly different from the definition of beneficence in the 1979 Belmont Report, which was the landmark report on research ethics published by the US National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report definition of beneficence amounts to an obligation to make 'efforts to secure [persons] well-being'.[4] The corollary to these rules are: (1) do not harm and (2) maximise possible benefits and minimise possible harms'. Crucially, Belmont appears to be relying on some sort of objective conception of harms and

[i]The exception to this rule is where a treatment is not clinically indicated, though even this might be seen as controversial by some commentators.

benefits, rather than a conception of harms and benefits that is relativised to the will and preferences of a patient. Autonomy and beneficence, in this sense, were once distinct principles, whereas in recent decades beneficence has come to be subsumed under the umbrella category of autonomy.[4] The good of the patient *just is* what the patient's considered preferences are.

Other bioethicists argue that autonomy is the most important ethical principle in bioethics and that it overrides other principles in cases of conflict. In a discussion of bioethical principles and their merits, Raanan Gillon argues that the principle of autonomy ought to be treated as 'first among equals'.[1] He outlines philosophical reasons why respect for autonomy should be of primary importance in medical ethics and applied ethics in general:

> autonomy…is what makes morality—any sort of morality—possible. For that reason alone autonomy—free will—is morally very precious and ought not merely to be respected, but its development encouraged and nurtured and the character traits or "habits of the heart" that tend to promote its exercise should indeed be regarded and extolled as virtues.[1]

Autonomy, in other words, is a necessary condition for morality of any kind, and we should ensure that—at the very minimum—we respect autonomy as it is expressed in the considered preferences of moral agents.

While there have been a range of arguments used for this position, it is clear that autonomy is seen as having a central and pre-eminent place among the ethical principles commonly invoked in bioethics discourse. More relevant for our purposes, however, is the relative lack of scrutiny afforded to the notion of autonomy given its centrality in bioethical theory and practice. In what follows, we will raise one significant theoretical concern about the notion of autonomy by exploring the notion of adaptive preferences.

WHAT IS AN ADAPTIVE PREFERENCE?

One underexplored way in which respect for patient autonomy can be difficult to conceptualise or apply in practice is in the case of some types of adaptive preferences. An adaptive preference is a preference which a person unconsciously forms in light of the options that they perceive themselves as having.[5] The classic example usually given is that of the fox who initially tried to reach a bunch of grapes. The fox fails to reach the grapes and sees that eating the grapes is no longer a feasible option for him. In response to this situation, the fox's preferences change—he does not want to eat the grapes anymore as he believes them to be sour.[6] Adaptive preferences are common in people's lives. In many cases these adaptations do not occur deliberately or consciously. Adaptive preferences may also persist long after a person's options have expanded to include a previously infeasible option. For example, the fox may continue to prefer not to eat the grapes even if his situation changes such that he is now able to reach the grapes.

Serene Khader has defined a subtype of adaptive preference which raises particular issues for medical ethics and moral philosophy. In Khader's terminology, an 'inappropriately adaptive preference' exists where people develop preferences which harm them in response to unjust social conditions.[7] Khader gives a number of examples of inappropriately adaptive preferences, such as the case of Lucy:

> Lucy's mother was a full-time caregiver and so was she. Lucy was raised to believe that women and men had different roles, and that women's moral duty was to take care of others. Her parents and gender ideology prevented her from training for a career outside of the home, and her husband would not have permitted her to work outside the home if she wanted to. Lucy was told her purpose in life was to nurture others—to make sure her husband was fed and listened to and had a clean house in the evenings, to make sure her children and grandchildren were safe and loved… Now, at the end of her life, Lucy is incapacitated such that she cannot care for anyone. She believes her life can have no meaning now and that it is thus time for her to accept palliative care and die.[8]

In this example, Lucy's preference to receive palliative care rather than medical treatment to extend her life is a direct result of unjust social arrangements (being prevented from taking up any educational or career option other than a full-time caregiver due to her gender). Her preferences have adapted to her feasible options, and harm her by causing her to refuse the life-extending treatment that she would otherwise choose.

Another case of an inappropriately adaptive preference could be a person who refuses life-sustaining, or life-extending treatment due to concern about developing a disability.[8] For example, suppose that a person lives in a society which severely discriminates against people with disabilities. As a result of their social conditions, the person believes that disability is shameful, people with disabilities are a burden on their families and that people with disabilities will inevitably be socially isolated and unable to access facilities to meet their needs. That person may develop the preference to refuse life-sustaining treatment rather than risk living with a disability. Their preference to die would then be a result of unjust social arrangements in which people with disabilities do not have access to dignified and reliable care, and are subjected to social exclusion and stigma.

A different example of an inappropriately adaptive preference in the context of medical decision-making would be the case of people who choose extensive cosmetic surgical interventions as a result of social prejudice or oppressive norms. It may be the case, for example, that a non-Caucasian person in a racist society wants to undergo surgery in order to look Caucasian. Alternatively, a female patient may request repeated cosmetic treatments to eliminate any trace of ageing, as she has internalised social prejudices that older women are ugly and not to be taken seriously. These cases should be distinguished from those in which a surgery would be refused on medical grounds, such as that of a patient diagnosed with body dysmorphic disorder, or cases in which surgery will be unlikely to achieve the desired result. This example of an inappropriately adaptive preference is one in which the desire for surgery is a preference caused by social factors rather than an identifiable individual mental illness.

It is important to note that whether a patient has an inappropriately adaptive preference is a very different question to whether that patient is 'competent' to make decisions about their own care. Some patients may lack decision-making capacity, due to dementia, psychosis, severe cognitive disability or other factors. In many cases, these patients can legally be judged to be incompetent to consent to or refuse medical care. However, in most cases where inappropriately adaptive preferences are the sole determining factor, the patient would clearly be legally competent to make choices about their own care.

It is easy to see why clinicians striving to practice in an ethical manner may be reluctant to comply with the patient's stated preferences in these cases. By acceding to inappropriately adaptive preferences, the clinician collaborates with the unjust social structures which gave rise to these preferences and thereby gives them some legitimacy and endorsement. By doing this, it is

possible that they may provide some benefit for the individual patient, but they do so by unjustly harming marginalised groups in society by reinforcing oppressive structures.

ARE INAPPROPRIATELY ADAPTIVE PREFERENCES AUTONOMOUS?

Some philosophers argue that inappropriately adaptive preferences should not be counted as autonomous, even if the person holding the preference might in general be an autonomous agent. This is because inappropriately adaptive preferences do not reflect the patient's authentic self. For example, Dale Dorsey argues that the reason that people care about preferences in moral and political philosophy is that they express what people value. However, an adaptive preference is different in that it does not genuinely reflect the subject's own point of view. As Dorsey puts it:

> We believe that the fact that someone values something has normative or evaluative consequences: it *makes* the thing valued valuable, or *pro tanto* worth pursuing. But sometimes, as in cases of adaptation, preferences do not genuinely express what someone values… In other words, adaptive preferences display a failure of autonomy—a failure to express what they really value… It is, as it were, putting on a kind of 'mask': adopting a preference or evaluative attitude that does not reflect me. That adaptive preferences are non-autonomous seems essential to the concept and function of preference-adaptation.[9]

In this way, inappropriately adaptive preferences do not have normative authority, because they are only a reflection of a person's social context and not of their true self.

Under this view, medical professionals who refuse to carry out actions based on inappropriately adaptive preferences are not in fact violating their patient's autonomy. If this was the case, there would need to be a way to distinguish between normal preferences (which are autonomous and thus have normative weight) and inappropriately adaptive preferences (which are non-autonomous and do not have normative weight), so that clinicians could know which patient preferences deserve respect.

However, as we will argue, it is very difficult to articulate a systematic and principled distinction between normal autonomous preferences and inappropriately adaptive preferences, especially if this distinction needs to be useful for clinicians in real-life situations. To illustrate this difficulty, we will outline five ways to characterise autonomous preferences: historical accounts, time-slice accounts, self-trust accounts, relational accounts and content-based accounts. We argue that none of them are useful to draw a clear dividing line between normal autonomous preferences and inappropriately adaptive preferences. This makes it hard to argue that inappropriately adaptive preferences are straightforwardly non-autonomous.

Historical accounts

One approach to this issue might be to develop a *historical* account of normal autonomous preferences, which focuses on the way these preferences were formed, rather than their content.[10–12]

One version of a historical account could suggest that what is distinctive about normal autonomous preferences is that people rationally and reflectively endorse the process by which they formed these preferences. For example, a person may rationally endorse their preference to live in Sydney because this preference was formed by a process of deliberation about the advantages and disadvantages of living in Sydney, given their career and family situation. Presumably this criterion would not be met in the case of inappropriately adaptive preferences, as people would not reflectively endorse the social conditions of deprivation, oppression and discrimination which result in inappropriately adaptive preferences.

However, this criterion seems to prove too much, and to wrongly designate many ordinary preferences as being non-autonomous and without normative weight. There are many ordinary preferences whose formation is somewhat arbitrary, and people do not rationally endorse the process by which they came to have these preferences. As Dorsey suggests, one example of this kind of preference is which football team a person supports.[9] Suppose that a person has a strong preference for the Canberra Raiders to win every game that they play. The person has developed this preference as a result of attending high school in Canberra and socialising with supporters of the Canberra Raiders. The person might not reflectively endorse the process by which they developed this preference, yet it would be counterintuitive to therefore say that their preference for the Canberra Raiders to win every game is an inappropriately adaptive preference or non-autonomous.

Another historical account could focus on whether a person made a decision to undertake a preference-formation process.[9] For example, suppose that a woman is initially neutral about parenting. She then makes a decision to adopt a child, and her experience of adoption gives her a deep attachment to the child. Her new preference in favour of parenting is autonomous and has normative weight because she made a decision to adopt a child, and the process which occurred as a result of her decision is the factor which has changed her preferences. It would seem that this criterion would not be met in the case of inappropriately adaptive preferences. It is unlikely that people made a decision to live in the deprived or oppressive conditions which lead to inappropriately adaptive preferences.

However, this account is no more plausible, as there are many situations in which intuitively autonomous preferences are a result of an unchosen process. For example, suppose that a person has a preference to eat Vegemite on toast every morning for breakfast. This preference is a result of growing up in a Vegemite toast-eating household in Australia. This person did not choose their parents or their family characteristics, yet it seems that their preference for eating Vegemite on toast is autonomous and not an inappropriately adaptive preference.

A supporter of these accounts may respond by arguing that this kind of test may be a viable way to distinguish between inappropriately adaptive preferences and normal autonomous preferences in a medical context, even if this does not always distinguish them in a non-medical context. They may argue that the nature of important medical decisions, such as a patient choosing from a range of recommended treatments, means that a process of deliberate reflective engagement with their preferences is necessary for the patient to choose autonomously.[ii]

However, there is no reason to believe that medical decisions are somehow special or are systematically different to other types of decisions. While medical decisions often have important implications for a person's well-being, so do other decisions, such as whether to marry or divorce, which career to pursue and where to live. Many preferences which affect these important

[ii]The authors thank an anonymous reviewer for this point.

life decisions—such as whether a person feels physical attraction to their partner and whether they feel 'at home' in a particular place—could be the result of an unchosen process which the person does not reflectively engage with. This does not thereby make these preferences non-autonomous.

Time-slice accounts

A different approach to this issue might be to develop an account of how normal autonomous preferences are structured in the present, without reference to the history of how these preferences were formed.[13] Like historical accounts, these accounts focus on the structure rather than the content of preferences.

One version of this might be to focus on whether a person rationally and reflectively endorses their current preference, while being informed of all relevant information. In this account, it is irrelevant how this preference was formed in the past, what is important is the person's rational endorsement of the preference in the present. It might appear that this criterion would exclude inappropriately adaptive preferences, as it would seem unlikely that people would rationally endorse preferences which harm them and reduce their well-being, especially if they are aware of the existence of other alternatives.

However, it is difficult to define what a rationally endorsed preference would be. In this context, a 'rationally endorsed preference' cannot just mean 'a preference which furthers a person's self-interest'. This would mean that all altruistic preferences would necessarily be defined as inappropriately adaptive preferences, which seems overly broad. Also, if self-interest is defined *subjectively*, then an inappropriately adaptive preference may actually further a person's self-interest. For example, a woman may gain more utility from boosting her self-image of herself as a self-sacrificing wife and mother, than she would from receiving medical treatment which would improve her health and extend her life.[14]

Another problem with this account is that there are many common aspects of preference formation and decision-making which could plausibly be described as irrational, such as availability bias, anchoring bias or confirmation bias. There is evidence that these biases are frequently present in everyday decision-making and the real-world practice of medicine itself.[15] It would be difficult to implement a system of medical ethics in which preferences were only respected as autonomous if they could be rationally endorsed. If this principle was consistently held, it is not clear how many patient (or clinician) preferences would be able to survive this kind of interrogation.

One possible test is to focus on the consistency across a person's preferences.[16] For example, suppose that a woman believed that in general women and men should be treated equally, but also that her health and quality of life are of much less importance than her husband's, and so she has a preference to forgo medical treatment. This arbitrary exception to general preferences may be an indicator that the lower value that she puts on her own health and well-being is an inappropriately adaptive preference. However, this is not a very helpful test in the real world. After all, suppose that the woman believed that women's health and quality of life is in general of much less importance than men's health and quality of life. This would not mean that her preference to forgo medical treatment would not be an inappropriately adaptive preference.

Another way to distinguish between inappropriately adaptive preferences and normal autonomous preferences might be whether the preference is consistent with the patient's current life plans, or their higher-order preferences.[13] For example, a patient may have an inappropriately adaptive preference to refuse medical treatment rather than live with a disability. This preference is inconsistent with their overall life plan—what the patient would really prefer is to live a life of normal length with all the social and employment opportunities accessed by people without disabilities. However, as Khader points out, it is questionable as to whether higher-order preferences are any less likely to be inappropriately adaptive.[14] It is entirely possible to have a life plan based on self-sacrifice and 'not being a burden' which is consistent with problematic medical preferences. Also, many autonomous preferences are inconsistent with overall life plans. If a person decides to eat ice cream for dinner rather than going to the gym, this may be inconsistent with their life plan of being fit and healthy, but is not thereby non-autonomous.[14]

Self-trust accounts

Some philosophers have argued that a particular core of self-trust or self-respect is a necessary precondition for autonomous preferences and decision-making.[17] For example, Trudy Govier suggests that: 'to lack general confidence in one's own ability to observe and interpret events, to remember and recount, to deliberate and act generally, is a handicap so serious as to threaten one's status as an individual moral agent'.[18] In this account, if a person lacks self-trust or self-respect, such that they do not see their values, beliefs and goals as having importance, and do not have confidence in their ability to perceive and act in the world, then they cannot have autonomous preferences.

It might be thought that the presence or absence of this self-trust could be used to distinguish between inappropriately adaptive preferences and normal autonomous preferences.[ii] Perhaps a person's experience of social oppression or discrimination may lead them to lack self-trust or self-respect, and then their lack of self-trust or self-respect may result in the development of inappropriately adaptive preferences.

However, the presence of self-trust is not useful for distinguishing between normal autonomous preferences and inappropriately adaptive preferences. A person can lack a great deal of self-trust, while still plausibly having autonomous preferences and not having inappropriately adaptive preferences. Imagine the case of a person who ignores warnings and repeatedly drives while drunk, eventually destroying their friend's car by driving into a tree. They then feel chastened by knowing that their pattern of reckless behaviour has caused significant damage and could easily have killed someone. After this incident they lack confidence about their values, deliberations and actions, and are unwilling to trust themselves. In this example, it seems implausible to conclude that the person has non-autonomous preferences or inappropriately adaptive preferences for as long as they lack self-confidence after the car crash. The lack of self-trust is better interpreted as a part of a process of moral self-improvement.

Even if a lack of self-trust did undermine a person's capacity to make autonomous decisions in general, it would not necessarily help us to identify the presence of specifically inappropriately adaptive preferences. There are many ways to undermine autonomy. For example, a person may not be able to make autonomous decisions due to suffering severe brain damage, but not have any inappropriately adaptive preferences.

Furthermore, a person can have inappropriately adaptive preferences even in the presence of self-trust. For example, an individual might have an inappropriately adaptive preference for extensive cosmetic surgery to avoid looking like a member of an oppressed ethnic group. This person may want the surgery not because they lack self-trust or confidence in their worth, but rather because they think that they are inherently better than

members of this ethnic group and do not want to have any resemblance to them. The presence of self-trust does not help distinguish between normal autonomous preferences and inappropriately adaptive preferences.

Relational accounts

Some philosophers have argued that whether a person has autonomous preferences depends on the nature of their relationships with other people. For example, Marina Oshana suggests an example of a woman who has no legal rights or opportunity for education, and no ability to travel, live or work independently. Every action depends on the permission of her husband or male relatives, and any show of independence is likely to result in punishment.[19] Oshana argues that this woman does not have autonomy, regardless of the content of her preferences, or the history or structure of how her preferences were formed. Oshana argues that autonomy is a property of a person's relationships with other people, and even if the woman genuinely has the same preferences as her husband or male relatives in every respect, she does not have the interpersonal standing to be autonomous. An account such as this may help distinguish inappropriately adaptive preferences from normal preferences—a clinician could identify whether a patient has inappropriately adaptive preferences by looking at the structure of the patient's interpersonal relationships.

However, the fact that a preference was formed in situations of social oppression does not necessarily mean that it is an inappropriately adaptive preference. Many people develop strong preferences to be respected as an equal person and to fight for justice as a result of having lived in an oppressive social situation. We would not want to characterise these preferences as non-autonomous and lacking normative weight.[20] In itself, being formed in a situation of social oppression or injustice is not enough to diagnose a particular preference as being inappropriately adaptive and non-autonomous.

Content-based accounts

A different approach is to say that some kinds of preferences are incompatible with autonomy because of their content. A person cannot have a properly autonomous preference unless that preference is objectively in their interest. For example, a preference to be subservient cannot be autonomous even if that preference was formed under ideal conditions and is reflectively endorsed, because living in subservience is not compatible with a person's objective well-being.[21] A content-based account of this sort would straightforwardly deny that inappropriately adaptive preferences are autonomous because of their problematic content.

However, this kind of account raises many problems. The self-determination which is characteristic of autonomy seems to be conceptually different from having a preference or choosing an action which is morally fitting or in a person's objective self-interest. Making an autonomous decision and making the right decision are two very different things. Furthermore, if a patient's preferences could be discounted as non-autonomous for having the 'wrong' content, then this forces clinicians to determine what is the 'right' preference content, and why preferences should have this content. There is likely to be widespread disagreement on this topic in practice.

It is also worth noting that this substantive conception of autonomy would be a radical redefinition from how autonomy has traditionally been discussed in the bioethics literature. For example, Beauchamp and Childress' account argues that choices are autonomous where they meet the minimal conditions of being intentional, adequately informed and not subject to external control.[22] Bioethicists typically have a 'thin' conception of autonomy[23] which can be very different from the term's usage in more abstract political philosophy discussions.

While there are many accounts of autonomy, this discussion aims to show that it is not a straightforward task to draw a dividing line between normal preferences (which are thought to be autonomous and having normative weight) and inappropriately adaptive preferences (which are not). It is difficult to articulate an account of autonomy which fits with people's moral intuitions, excludes inappropriately adaptive preferences and is straightforward to use in real-life clinical situations.

IMPLICATIONS FOR THE THEORY AND PRACTICE OF MEDICAL ETHICS

Theoretical issues with autonomy

If autonomy is to do the moral work that it is often expected to do in bioethics, it would certainly be desirable that we had an account of autonomy that provided us with a clear means of distinguishing normal preferences from inappropriately adaptive preferences. However, it is difficult to have a working definition of autonomy according to which all normal preferences are distinguished from all inappropriately adaptive preferences. This seems to us a significant and under-discussed theoretical issue inherent in the concept of autonomy as it is often deployed in bioethics discourse.

This, in turn, raises questions about patients' autonomous preferences in medicine. It might be thought that—provided a patient is competent and adequately informed in their decision-making—medical professionals should always defer to patient choices where the patient has expressed a firm and definite treatment preference. The existence of inappropriately adaptive preferences, however, suggests that clinicians may have reason to be more hesitant about a certain subset of patient preferences. We should exercise caution with respect to patient preferences that are shaped by manifestly prejudicial social beliefs and that will likely perpetuate social injustice in one form or another. We are not saying that some patient preferences should be ignored or rejected. It is appropriate, however, to think about ways in which healthcare professionals might be more aware of the moral nuances surrounding patient choices based on inappropriately adaptive preferences. To this end, the next subsection will consider how healthcare professionals can engage with social prejudice when prejudice manifests itself in the preferences of patients.

Responding to inappropriately adaptive preferences in medical education and practice

During their career, clinicians are likely to encounter patients who have inappropriately adaptive preferences. It is important, therefore, to consider suggestions for how inappropriately adaptive preferences could be addressed in practice, and what implications these preferences could have for clinical ethics education.

One suggestion could be to make greater use of therapies which have been developed to deal with problematic cultural assumptions that inform patient preferences. One example is dignity therapy—a psychotherapy to relieve psychological and existential distress in patients at the end of life.[24] In a palliative context, for example, dignity therapy may be used to help address a patient's sense that they have lost any semblance of dignity. Importantly, a patient's desire for death or experience of demoralisation at the end of life may be influenced by cultural assumptions that they have internalised about the nature of dignity and

its relationship to independence and being in control of one's life. Someone may, for example, have a distorted conception of dignity according to which one cannot have dignity without dominion over one's own life, including the physical ability to care for oneself and to independently navigate the challenges of life.[25] This may lead one to feelings of self-disgust and loathing when one is forced by terminal illness to depend on the support of family and healthcare staff. Dignity therapy, however, can help patients to come to terms with the experience of dependence and a lack of control by helping patients to focus on their life history, achievements, significant relationships with loved ones and the attainment of closure through creating a permanent record of one's life and experiences. At the very least, it can alleviate anxiety surrounding death and dying that is often experienced by terminally ill patients.

To be clear, our claim is not that the preferences of patients ought to be dismissed when they are informed by what might reasonably be deemed *problematic cultural assumptions*. One could argue that such an approach would be paternalistic. Rather, we are simply arguing that healthcare staff, when encountering patients who express these preferences, ought to consider whether therapeutic modalities (such as dignity therapy) may be appropriate to recommend to patients given the complex cultural factors that underpin phenomena such as the experience of pain and death anxiety.[26]

Another suggestion would be to make changes to medical ethics education, so that clinicians are made aware of the complexity of respecting autonomy in clinical practice. Respect for patients' autonomy may seem a straightforward principle in a textbook, but the case of inappropriately adaptive preferences indicates that understanding and applying it in the real world can be difficult and ambiguous. It may be easier to say that respect for autonomy is the pre-eminent principle of medical ethics, rather than to deal with a plurality of values which may seem to compete with each other. However, the principle of respect for autonomy does not have the conceptual resources to override other values and alone be the guiding principle for ethical clinical practice.

Clinical ethics education in medical schools should be improved to equip students with the tools necessary to handle the moral challenges arising in contemporary medical practice.[27] Medical students and clinicians receiving ethics education should be able to critically evaluate how a variety of ethical principles relate to each other and ought to be applied in practice. At the very least, any attempt to introduce students to Beauchamp and Childress's four principles ought to be complemented with an exploration of how one balances these principles with each other, rather than adopting an approach whereby patient preferences override all other moral considerations. It is likely to be helpful to introduce educational modules in medical schools which illustrate the challenges which may arise due to inappropriately adaptive preferences. It may also be useful to have experienced clinicians address students about how they negotiate clinical dilemmas involving inappropriately adaptive preferences. Surely there is some degree of practical wisdom required to navigate these and other ethical challenges that arise in the course of clinical work.

CONCLUSION

In this paper we have discussed the notion of inappropriately adaptive preferences and have considered the implications of this concept for the theory and practice of respect for autonomy in medicine. We argued that it is difficult in practice to draw a dividing line between normal autonomous patient preferences which should be respected, and non-autonomous inappropriately adaptive preferences which carry much less normative weight. This makes it hard to argue that inappropriate adaptive preferences are straightforwardly non-autonomous.

We then considered how a robust understanding of inappropriately adaptive preferences might inform the way in which autonomy is theorised in contemporary bioethics. We argued that a certain subset of arguably autonomous preferences may in fact conflict with obligations to fight social injustice and prejudice in broader society. We discussed what this might mean for the practice of medicine and medical education. More research exploring the intersection between inappropriately adaptive preferences and autonomy would be fitting, particularly in light of the status of autonomy in contemporary bioethics as a principle that is 'first among equals'.

Twitter Xavier Symons @xaviersymons

Contributors SP had the original idea for this paper and drafted the sections on adaptive preferences. XS drafted the sections on the medical ethics literature and medical education and practice. Both SP and XS approved the final draft. SP is the guarantor.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study does not involve human participants.

Provenance and peer review Not commissioned; externally peer reviewed.

There are no data in this work.

REFERENCES

1 Gillon R. Ethics needs principles--four can encompass the rest--and respect for autonomy should be "first among equals". *J Med Ethics* 2003;29(5):307–12.
2 Tauber AI. Sick autonomy. *Perspect Biol Med* 2003;46(4):484–95.
3 Beauchamp T. Who deserves autonomy, and whose autonomy deserves respect? In: Taylor JS, ed. *Personal autonomy: new essays on personal autonomy and its role in contemporary moral philosophy*. London: Cambridge University Press, 2005.
4 Lysaught MT. Respect: or, how respect for persons became respect for autonomy. *J Med Philos* 2004;29(6):665–80.
5 Stoljar N. Autonomy and adaptive preference formation. In: Veltman A, Piper M, eds. *Autonomy, oppression, and gender*. Oxford University Press, 2014.
6 Bruckner DW. Adaptive preferences, autonomy, and extended lives. In: *Adaptation and autonomy: adaptive preferences in enhancing and ending life*. Berlin, Heidelberg: Springer, 2013: 7–26.
7 Khader SJ. *Adaptive preferences and women's empowerment*. New York: Oxford University Press, 2011.
8 Khader SJ. Gendered Adaptive Preferences, Autonomy, and End of Life Decisions. In: *Adaptation and autonomy: adaptive preferences in enhancing and ending life*. Berlin, Heidelberg: Springer, 2013: 81–100.
9 Dorsey D. Adaptation, Autonomy, and Authority. In: *Adaptation and autonomy: adaptive preferences in enhancing and ending life*. Berlin, Heidelberg: Springer, 2013: 27–47.
10 Christman J. *The politics of persons: individual autonomy and socio-historical selves*. Cambridge: Cambridge University Press, 2009.
11 Dworkin G. *The theory and practice of autonomy*. Cambridge: Cambridge University Press, 1988.
12 Friedman M. *Autonomy, gender, politics*. New York: Oxford University Press, 2003.
13 Frankfurt HG. *The importance of what we care about: philosophical essays*. New York: Cambridge University Press, 1988.
14 Khader SJ. Adaptive preferences and procedural autonomy. *J Human Dev Capabil* 2009;10(2):169–87.
15 Saposnik G, Redelmeier D, Ruff CC, *et al*. Cognitive biases associated with medical decisions: a systematic review. *BMC Med Inform Decis Mak* 2016;16(1):1–4.
16 Mitchell P. Adaptive preferences, adapted preferences. *Mind* 2018;127(508):1003–25.
17 McLeod C. *Self-trust and reproductive autonomy*. Cambridge: MIT Press, 2002.
18 Govier T. Self-Trust, autonomy, and self-esteem. *Hypatia* 1993;8(1):99–120.
19 Oshana M. *Personal autonomy in society*. London: Routledge, 2016.
20 Stoljar N. Feminist Perspectives on Autonomy. In: Zalta EN, ed. *The Stanford encyclopedia of philosophy*, 2018. https://plato.stanford.edu/archives/win2018/entries/feminism-autonomy/
21 Babbitt S. Feminism and Objective Interests: The Role of Transformation Experiences in Rational Deliberation. In: Meyers DT, ed. *Feminist social thought: a reader*. New York: Routledge, 1993.

22 Beauchamp TL, Childress JF. *Principles of biomedical ethics*. New York: Oxford University Press, 2012.
23 Reis-Dennis S. Understanding autonomy: an urgent intervention. *J Law Biosci* 2020;7(1):lsaa037.
24 Martínez M, Arantzamendi M, Belar A, *et al*. 'Dignity therapy', a promising intervention in palliative care: a comprehensive systematic literature review. *Palliat Med* 2017;31(6):492–509.
25 Varelius J. Refusing Life-Saving Treatment, Adaptive Preferences, and Autonomy. In: *Adaptation and autonomy: adaptive preferences in enhancing and ending life*. Berlin, Heidelberg: Springer, 2013: 183–97.
26 Mehta A, Chan LS. Understanding of the concept of" total pain": a prerequisite for pain control. *J Hosp Palliat Nurs* 2008;10(1):26–32.
27 Carrese JA, Malek J, Watson K, *et al*. The essential role of medical ethics education in achieving professionalism: the Romanell report. *Acad Med* 2015;90(6):744–52.