

Knowing and Not-knowing for your own good: The Limits of Epistemic Paternalism

Forthcoming in the Journal of Applied Philosophy

Emma C. Bullock

bullocke@ceu.edu

Abstract: Epistemic paternalism is the thesis that a paternalistic interference with an individual's inquiry is justified when it is likely to bring about an epistemic improvement in her. In this paper I claim that in order to motivate epistemic paternalism we must first account for the value of epistemic improvements. I propose that the epistemic paternalist has two options: either epistemic improvements are valuable because they contribute to wellbeing, or they are epistemically valuable. I will argue that these options constitute the foundations of a dilemma: either epistemic paternalism collapses into general paternalism, or a distinctive project of justified epistemic paternalism is implausible.

Keywords: Epistemic paternalism; Paternalism; Epistemic Value; Disinformation

Introduction

Roughly speaking, a paternalistic interference is one in which an individual's choices or actions are interfered with for that individual's own good.¹ Part of determining whether a paternalistic interference is justified involves identifying what counts as a benefit to the individual interfered with. Whilst much of the literature on paternalism has focused on identifying the prudential pay-offs of paternalistic interferences,² a burgeoning discussion proposes that we could think about the benefit in epistemic terms.³ Epistemic paternalism is motivated by the thought that the individual is more likely to acquire, retain and make good use of true beliefs (according to the veritist),⁴ or increase understanding (according to the virtue epistemologist),⁵ should an interference take place. Discussions of justified epistemic paternalism thus equate the practice with the project of 'epistemic amelioration.'⁶

In this paper I argue that in order to motivate this project the epistemic paternalist needs to explain why epistemic improvements are valuable (section 2). I propose that there are two options available to the epistemic paternalist: either the value of epistemic improvements is explained in terms of their contribution to wellbeing or it is explained in terms of epistemic value. I argue that the choice to be made between these options is the locus of a dilemma. First, an appeal to the eudaimonic value of epistemic improvement renders the position of epistemic paternalism indistinct from general paternalism (section 3). The epistemic paternalist must thereby concede that paternalistic interferences that worsen an individual's epistemic situation are sometimes justified. Second, an appeal to epistemic value makes it implausible that epistemic paternalism is ever justified (section 4). Therefore, either there is no moral distinction between epistemic paternalism and paternalism in general, or if it is distinctive, a project of justified epistemic paternalism is intuitively implausible.⁷ The onus is on the epistemic paternalist to defend their counter-intuition.

1. The Nature of Epistemic Paternalism

According to a recent definition, a practice is epistemically paternalistic iff:

1. It does not consult those interfered with on the issue of whether they should be interfered with in the relevant manner (*the non-consultation condition*)
2. It interferes with the freedom of inquirers to conduct inquiry in whatever way they see fit (*the interference condition*)
3. It interferes – exclusively or not – for the purpose of making those interfered with epistemically better off (*the improvement condition*).⁸

The first condition requires that the interference bypasses the agency of the individual interfered with. This is a condition that applies generally to paternalistic interferences; in order for an interference to count as *paternalistic* (as opposed to a consensual interaction) the interference must fail to consult the individual as to whether she wants to be interfered with in the relevant manner.⁹

The second defining condition of epistemic paternalism is that it interferes with the freedom of inquirers to conduct inquiry in whatever way they see fit. There are a number of ways in which an individual's inquiry can be paternalistically interfered with.¹⁰ First, information can be paternalistically disclosed to the individual. Health education measures could count as an example of this, should individuals prefer that information about the ways in which they are failing to live a healthy lifestyle are not disclosed to them.¹¹ The practice of divulging unwanted information has also been discussed at length in terms of a patient's 'right not to know' about her medical condition.¹² Those who argue that such information should be mandatorily disclosed are defending a kind of epistemic paternalism.¹³

Second, an individual's inquiry can be paternalistically interfered with by withholding information from her. Withholding information from a jury about the past crimes of a defendant, for instance, is thought to improve the jury's ability to deliberate over the defendant's innocence.¹⁴ Other familiar cases of withholding information include the restriction of curricula in educational settings and the regulation of commercial advertising.¹⁵ These restrictions are partly rationalized on the basis that individuals will be in a better epistemic situation in the absence of false or confusing information.¹⁶

Third, an individual's inquiry can be interfered with through the use of deception, examples of which can also be taken from educational settings. For instance, a tutor might deceive her

students into believing that there will be a surprise exam at some point throughout their university course, in order to motivate them to revise more frequently (and thereby improve the course of their inquiry).¹⁷ The educational tool of ‘lying-to-children’ can also be viewed as a kind of deceptive interference with an individual’s inquiry. A ‘lie-to-children’ involves teaching false or incomplete theories to students in order to facilitate a better understanding of more complex theories. For instance, a student might first be taught that Newtonian mechanics provides a complete account of the laws of motion, in order to make it easier for them to learn quantum mechanics.¹⁸

Fourth, an individual’s inquiry could also be interfered with through coercive measures. Whilst it is plausible to think that it is impossible to directly coerce beliefs,¹⁹ there are a number of indirect ways in which an individual’s beliefs or understanding can be coerced. One way would be through physically manipulating the brain and directly implanting beliefs into the individual. Another way of coercing a change in an individual’s inquiry would be via a physical threat. Imagine, for instance, that you hold a gun to my head and threaten to kill me unless I believe that the current prime minister is a clone. Whilst this is not enough for me to form the belief (despite my desperate claims otherwise), it is enough to make me behave in such a way as to acquire the belief; the threat might motivate me to read your literature on the matter, or even to take a pill guaranteed to induce the belief.²⁰

It is worth noting that each of these ways of interfering with an individual’s inquiry are available to the general paternalist. Indeed, some of the practices defended by ‘nudge’ paternalists involve reframing information in order to bring about (non-epistemic) welfare improvements in the individual.²¹ For instance, informing individuals that they are at an increased risk of developing skin or breast cancer if they fail to engage in self-examinations (as opposed to informing them of

the reduced risk should they carry out self-examinations) is thought to increase the chances that they will do so.²² Motivating individuals to carry out self-examinations in order to thereby reduce the risk of developing untreatable cancers can thus be facilitated by interfering with their inquiry.

In the case of general paternalism the individual's inquiry is interfered with in order to bring about some non-epistemic benefit (such as an improvement to her health). The distinctive feature of epistemic paternalism thus lies in condition three, according to which the purpose of an epistemically paternalistic interference is to make those interfered with *epistemically better off*. To date, the nature of an epistemic improvement has been drawn along veritist lines (an individual's epistemic situation is improved should she be more likely to acquire, retain and make good use of true beliefs)²³ or in terms of fostering better understanding.²⁴ I will not endorse either view here since the thrust of my argument does not depend on the account adopted.

1.1 Types of Epistemic Paternalism

Following the literature on general paternalism it is possible to make distinctions between types of epistemic paternalism.²⁵ The first distinction can be made in terms of whose liberty is restricted. With *direct* epistemic paternalism the individual interfered with is identical to the individual that benefits. Epistemic paternalistic interferences that interfere with an individual in order to benefit some other person(s) are properly thought of as *indirect*.²⁶ A clear example of direct epistemic paternalism is divulging unwanted medical information to a patient; it is the patient's inquiry that is interfered with and the patient that is supposed to benefit. Deceiving an undergraduate student into believing there will be a surprise exam could count as a case of indirect epistemic paternalism if the interference is carried out for the epistemic benefits this will have on her peers: her contributions in class are better informed and she is able to help them with their assignments.

A second distinction between types of epistemic paternalism can be drawn in terms of how the interference is motivated. *Mixed* epistemic paternalistic interferences are motivated by additional non-epistemic reasons. An *unmixed* epistemic paternalistic interference is motivated only for the reason that the interference will make the individual epistemically better off.²⁷ A good example of a mixed epistemic paternalistic interference is the case of withholding information from a jury. One reason this information is withheld is to bring about an epistemic improvement in each juror, but a further reason (and perhaps the main reason) is in order to facilitate justice – which is of benefit to society (including the jury). Note however that this could equally be a case of unmixed epistemic paternalism should the epistemic paternalist only be interested in bringing about epistemic improvements in the jury.

The final distinction rests on the competency of the individual interfered with. A *soft* epistemic paternalistic interference takes place with an individual who is judged as incapable of making decisions about their inquiry. A child's education perhaps falls into this category. A *hard* epistemic paternalistic interference takes place with individuals who are judged to be worthy of having their decisions about their inquiry respected.

The distinctions between direct/indirect, mixed/unmixed and soft/hard can be combined with each other in interesting ways. In the following my focus is on hard direct epistemic paternalism: specifically, on interferences with a competent individual's inquiry for that individual's own epistemic good. Whilst the examples used in the remainder of the paper could be interpreted in a number of ways, they should be read as examples of interferences with an individual for that individual's own epistemic good.

2. Epistemic Value

Paternalistically interfering with an individual on the basis that this will lead to an epistemic improvement in her is a controversial practice. For instance, some individuals would prefer not to be confronted with health warning messages on cigarette packaging if they *want* to be addicted to tobacco. Such an individual might want to have: "...a classical 'rockstar' lifestyle, which implies to live in the present, to not care about the future, and to indulge in all the pleasures associated with this lifestyle."²⁸ Equally, patients don't always want to be informed about their medical condition, claiming that divulging this information violates their autonomy.²⁹

In order to address such controversies the epistemic paternalist needs to both motivate and justify their position. To motivate the position the epistemic paternalist needs to provide an explanation of *why* epistemic improvements are valuable. To justify the position they need to explain *when* epistemic paternalistic interferences are justified. To date discussions have focused on when epistemic paternalism is justified, identifying the constraints that need to be in place in order for epistemic paternalism to be practiced.³⁰ This discussion is only worthwhile, however, insofar as the project is properly motivated. This is because if one is unconvinced that we have reason to epistemically paternalistically interfere at all, then one will not care *when* epistemic paternalistic interferences are supposedly justified.

The general paternalist motivates her position by arguing that paternalistic interferences are likely to bring about an improvement in wellbeing. Not much more is needed than this to motivate the position: wellbeing, flourishing or living a good life are all things that are intuitively valuable. (Indeed, the dispute between paternalists and anti-paternalists tends to focus on whether and the degree to which wellbeing value outweighs personal autonomy rather than on whether wellbeing is of any value at all). In order to motivate epistemic paternalism a similar move needs to be made. This will amount to appealing to the value of epistemic improvement. Unlike the

general paternalist, however, the epistemic paternalist needs to further explain what this value is, for whilst the value of wellbeing is intuitive it is fairly difficult to get a grip on what is meant by the value of epistemic improvement without further explanation. (For this reason, unlike with general paternalism, a very real objection to epistemic paternalism is that epistemic improvements are not valuable).

There are two broad routes that the epistemic paternalist can take to explain the value of epistemic improvements. The first way identifies it with eudaimonic value, i.e. value vis-à-vis well-being.³¹ A second route is to identify it with epistemic value. On such accounts: “[t]he ignorant would always be ‘epistemically better off’ were they knowledgeable, even if not better off.”³² From here on I will refer to epistemic paternalism that adopts the former explanation as “Eudaimonic Epistemic Paternalism” and the latter as “Strict Epistemic Paternalism.”

3. Eudaimonic Epistemic Paternalism

The eudaimonic epistemic paternalist motivates her position by appealing to the value of wellbeing. There are two ways in which this can be further cashed out. First, epistemic improvements can be thought of as instrumentally valuable to wellbeing. On this account, epistemic goods are valuable because they make it more likely that the individual will be better off in terms of her wellbeing. Second, epistemic improvements can be construed as valuable because they are constitutive of wellbeing. On either theory, if we grant that the value of wellbeing is a good ground for justifying paternalism in general, then it looks as though we have grounds for motivating eudaimonic epistemic paternalism.

Problematically, the eudaimonic epistemic paternalist’s neat answer to the motivation question comes at the cost of losing the position’s moral distinctiveness. To bring this out it is

worth recalling the defining conditions of epistemic paternalism (Section 1). Recall first that general paternalistic interferences can meet both the non-consultation condition and the interference condition, since they can also involve interferences with an individual's inquiry. I argued that the distinctiveness of epistemic paternalism is instead that the purpose of the interference is to make those interfered with epistemically better off (the improvement condition). However, if the value of an epistemic improvement lies in its contribution to wellbeing then these interferences fall under general paternalism: bringing about epistemic improvements is just one of the ways to achieve the general paternalist's goal. Since the general paternalist can meet all three defining conditions of eudaimonic epistemic paternalism the two positions are indistinct.

The epistemic paternalist might respond that what makes the project distinctive is that paternalistic interferences are *only* justified for the sake of bringing about epistemic improvements. This position, however, is absurd. To bring this out, consider other possible paternalisms making a similar claim. Examples could include, for instance, 'dental health' paternalism, 'relationship' paternalism or 'good night's sleep' paternalism. These paternalisms are absurd insofar as they claim to be exclusive: i.e., that the only times we can paternalistically interfere are for the sake of improving dental health or fostering relationships or inducing sleep. The reason we value dental health, good relationships and a good night's sleep is because of their contribution to wellbeing; to limit justified paternalistic interferences to any one of these goods is therefore arbitrary. The same holds for eudaimonic epistemic paternalism.³³

The collapse of epistemic paternalism into general paternalism has significant repercussions for the project of epistemic amelioration. This is because what is epistemically good for us can come apart from what is eudaimonically good for us.³⁴ In the following I present a number of cases in which the eudaimonic epistemic paternalist must concede that the individual

might be better off for being made epistemically worse off. This looks bad for the project of epistemic amelioration, since it suggests that we are potentially justified in manipulating and coercing individuals in order that they are made epistemically worse off.

3.1 The War on Drugs

One example of paternalism that purposefully aims to bring about an epistemic decline in the individual is illustrated by the 'war on drugs'. Government prohibitionists in the US have been known to 'routinely lie' and 'wilfully mislead' the public about the harms of drug consumption.³⁵ Wilful disinformation about the harms of drug consumption, such as the suggestion that marijuana creates 'crazed killers', and that crack cocaine is 'instantly addictive' has been criticized for the very reason that it makes the citizenry epistemically worse off, and is thereby inconsistent with 'enduring education.'³⁶

We can see, however, that the dissemination of disinformation about the harmful effects of drug consumption could be justified on paternalistic grounds under the eudaimonic account of epistemic value. Specifically, we might have good reason to suppose that bringing about an epistemic decline in the individual will improve their wellbeing: citizens are likely to be less tempted to try drugs if they falsely believe that they are extraordinarily harmful, and so more likely to lead a flourishing drug-free life. There is some reason to think that disinformation is unjustified in this case because it causes more harm than good.³⁷ What is important to see is that if the practice did in fact improve the individual's wellbeing, then the eudaimonic epistemic paternalist cannot claim that the policy is unjustified because it brings about an epistemic decline.³⁸

3.2 Health

A second example of paternalism that depends upon bringing about an epistemic decline in the individual can be taken from the use of placebos in both medical research and therapeutic practice. A placebic treatment is one that relies on pretence in order to bring about a therapeutic improvement in the patient. This might involve, for instance, prescribing pharmacologically inert sugar capsules to a patient on the basis that they will alleviate the patient's symptoms.³⁹ Other examples are the use of saline injections in order to produce an analgesic effect in the patient and the use of 'sham' surgeries.⁴⁰

Significantly, in order for placebic treatments to be effective, the patient must be deceived about the efficacy of the treatment being received.⁴¹ The deception is directly used in order to bring about a health improvement in the patient. The practice of deceiving patients through the administration of placebos could thus be justified should it bring about an improvement in the patient's overall wellbeing. The eudaimonic epistemic paternalist is unable to rule out the justifiability of such interferences on their own terms.

Sometimes true beliefs can be all-things-considered bad for an individual's health.⁴² Informing a patient of the potential side effects of a prescribed drug can have a 'nocebo' effect, whereby the information itself produces the side effects independently of the pharmacological properties of the drug.⁴³ In some contexts there is evidence that informed patients suffer more anxiety and pain and have more postsurgical complications than their uninformed counterparts.⁴⁴ Paternalistically withholding true information from patients could thus be better for them all-things-considered because patients would suffer more pain with knowledge than without it. There is also some evidence that positive illusions about one's medical condition can enhance health.⁴⁵ Not only might the eudaimonic epistemic paternalist be justified in paternalistically withholding

true information, but in some cases justified in fostering illusions or false information for the sake of bringing about an improvement in an individual's health.⁴⁶

3.3 Cognitive Bias

The literature on cognitive bias has been particularly influential for the project of motivating paternalism.⁴⁷ Unsurprisingly, the problem of cognitive bias has also been used to motivate epistemic paternalism.⁴⁸ Somewhat ironically for the eudaimonic epistemic paternalist, there is some reason to think that cognitive biases can be all-things-considered good for us. Consider, for instance, self-enhancement bias in which one favours an overly positive conception of oneself. Whilst this is epistemically bad for us, it has been suggested that an overly positive self-evaluation can help avoid depression and promote high self-esteem.⁴⁹ Inflated self-perception is also associated with increased chances of success and achievement.⁵⁰ Rather than seeking to free individuals from their cognitive biases, the eudaimonic epistemic paternalist thus has reason to foster some cognitive biases, to discourage individuals from trying to overcome them or even to paternalistically lie to individuals about their existence.

In sum, by motivating epistemic paternalism with an appeal to wellbeing, epistemic paternalism collapses into general paternalism. This means that the epistemic paternalist must concede that in some cases we are justified in paternalistically manipulating or coercing individuals so that they become epistemically worse off. Whilst the epistemic paternalist might stipulate that interferences that bring about an epistemic deterioration are unjustified they cannot do this in a non-arbitrary way.

4. Strict Epistemic Paternalism

Strict epistemic paternalism saves the moral distinctiveness of epistemic paternalism by identifying the value of epistemic improvements independently of wellbeing. I will argue in the following, however, that it is unclear that epistemic value is weighty enough to meet the justificatory constraints on its implementation.

4.1 Justification

Two suggested conditions for justifying an epistemically paternalistic interference are as follows:

- a. **Epistemic motivation:** we have good reason to think that individual will be made epistemically better off⁵¹
- b. **Balancing goods:** we have no strong reason to suppose that the inquirer will be made all-things-considered worse off.⁵²

The first constraint on justified epistemic paternalism is that we have good reason to believe that the interference will bring about an epistemic improvement in the individual interfered with. This means that interfering with an individual on the basis of a crazy or false belief that it will bring about an epistemic improvement is unjustified. Consider the following example:

Malevolent force: I believe that a malevolent force is negating the information I divulge to other people. Consequently, I believe that if I tell you a truth: i.e. that smoking is bad for your health, then the malevolent force will make you believe a falsehood: that smoking

is *not* bad for your health. Correspondingly, I believe that if I tell you that smoking is not bad for your health, then the malevolent force will make you believe the truth that smoking is bad for your health. On the basis of this crazy belief, I tell you a falsehood because I think that this is the best means of bringing about an epistemic improvement in you.

The above example is not a case of justified epistemic paternalism because I do not have any good reason for supposing that you will go through an epistemic improvement. The reason I have is a poor one because there is no malevolent force interfering with your inquiry and I have no good reason to think otherwise.

The balancing goods condition is motivated by the thought that we might have very strong reasons against interfering even if the interference brings about an epistemic improvement. Ahlstrom-Vij expresses the worry as follows:

... we can imagine a government with such complete control over government organs, news media outlets, educational institutions and so on that it is able to mandate the use of nothing but the most reliable scientific methods in virtually every domain of life, and remove from public consumption any misleading or biasing information. It might be that this would do epistemic wonders for the citizenry, and the government might to that extent be motivated to exercise the relevant form of control on epistemic grounds. But...of course...we might have independent reason – indeed, possibly very *strong* independent reason – to hesitate about a government exercising such great control, even given the prospects for significant epistemic benefits. That is, we might have non-epistemic reasons against having a government exercise epistemic paternalism on such a great scale.⁵³

The thought here is that even if an interference is likely to bring about an epistemic improvement, we should not disregard the effect that this might have on the agent's overall wellbeing, her moral responsibility, or indeed on matters of justice.⁵⁴ One way in which non-epistemic harms can be taken into account is by balancing the epistemic benefits against them, hence the balancing goods condition (b).

Ahlstrom-Vij (2013) criticizes the approach of 'balancing goods' on the basis that there are situations in which it is unclear how epistemic and non-epistemic factors can be weighed against one another, making it difficult to know when an epistemically paternalistic interference is justified. Ahlstrom-Vij thus drops the balancing goods requirement in favour of the *alignment condition* according to which the epistemic reasons we have for the interference need to be aligned with our non-epistemic reasons. In order to be aligned, our non-epistemic reasons must either be (a) additional reasons for interfering, or (b) neutral (which amounts to saying that we have no non-epistemic reasons against interfering).⁵⁵

The alignment condition is, however, overly restrictive. This is because there are cases in which we can have a weak non-epistemic reason against interfering, but nonetheless think that the epistemic paternalistic interference is justified. Take the following example:

Daphne is considering motherhood. She knows that she may have a chance as high as one in four of having a child with Tay-Sachs disease. Her doctor has the results of her genetic test, which indicate that she does not carry the genetic trait, thereby ruling out the possibility that any of her children will be born with the disease. Since agreeing to carry out the test, Daphne has become pregnant and has decided that she no longer wants to know

whether she is at risk, preferring to adopt an attitude of responsibility to her child whatever the outcome. During the pregnancy, Daphne prepares for what she believes to be the possible outcome that her child will be born with the disease and her anxiety increases. In reality, she is worrying for no reason.⁵⁶

Given that informing Daphne that she does not carry the genetic trait would violate Daphne's reproductive autonomy we have a non-epistemic reason not to interfere and so the alignment condition tells us that the revelation of this information would be unjustified. Intuitively, however, the epistemic and non-epistemic benefits of divulging this information to Daphne most likely outweigh the non-epistemic harms.

Ahlstrom-Vij concedes that there may be situations in which an epistemically paternalistic interference fails to satisfy the alignment condition but would nevertheless be justified on balancing grounds. He argues, however, that these situations are not the norm.⁵⁷ The failure of alignment is, however, pervasive for direct epistemic paternalism; there will always be a non-epistemic reason against interfering, namely, the harm it would cause to the individual's personal sovereignty. This issue has been largely ignored because the permissibility of epistemic paternalism has been measured against the harm it would bring to the inquirer's epistemic autonomy.⁵⁸ According to accounts of epistemic autonomy, the epistemically autonomous individual relies on no one else for her knowledge and: "...accepts only what she has found out for herself, relying only on her own cognitive faculties and investigative and inferential powers."⁵⁹ The thought that epistemic autonomy is an ideal to strive towards is correctly criticized. For one thing, the ideal is not feasible.⁶⁰ Second, the ideal is unattractive, and so it hard to see why it is valuable to be epistemically autonomous.⁶¹ Thus, the argument runs, since epistemic autonomy is

of no (or very little) value, it does not count against a justified epistemic paternalistic interference.⁶²

The objection to epistemic autonomy is strong. However, the focus on epistemic autonomy has overlooked an important point: that the permissibility of epistemic paternalism still needs to be measured against our reasons for respecting personal sovereignty in the sense that the individual has a claim on whether or not she is interfered with. According to the anti-paternalist, an individual who can act autonomously and bears an autonomy right has personal sovereignty over: "...all those decisions that are 'self-regarding,' that is which primarily and directly affect only the interests of the decision maker."⁶³ They further claim that an individual's personal sovereignty is of such value that we morally ought never disrespect or interfere with their autonomy out of concern for her wellbeing. This anti-paternalistic point can be straightforwardly expanded as a limit on concerns for her epistemic improvement, as it is with respect to interferences that are carried out for the sake of improving moral character or religious conviction.⁶⁴

One way of salvaging the alignment condition would be to deny that we have reason to respect personal sovereignty. This, however, would be a radical view, comparable with an anti-paternalist denying that there is any value to wellbeing. A more moderate response would be to claim that personal sovereignty is only *pro tanto* or *prima facie* valuable. However, if personal sovereignty is only *pro tanto* valuable, then in order to find out how far that value reaches in a particular case it will still need to be balanced against considerations of epistemic improvement. Likewise, to figure out whether or not personal sovereignty has actual value or merely the appearance of value in a particular context we would need to balance it against the other values we care about. The alignment condition therefore fails to give an account of when epistemically paternalistic interferences are justified because considerations of personal sovereignty will always

give us a potential reason against interfering. The epistemic paternalist should thus hold onto the balancing condition.

4.2 The weight of epistemic value

Unlike eudaimonic epistemic paternalism, strict epistemic paternalism is a morally distinctive project. So far I have argued that in order to be justified a strict epistemically paternalistic interference needs to meet the conditions of (a) epistemic motivation and (b) balancing goods. The question that remains is whether the justificatory conditions can be met, specifically: whether epistemic value can outweigh considerations of personal sovereignty.

As noted above, epistemic paternalists have tried to justify their project by balancing epistemic improvements against the harms to epistemic autonomy. But now for the first time we can clearly compare autonomy proper with epistemic value, it's intuitive that the project is unjustified. Suppose, for example, that I play a series of physics lectures to you whilst you are sleeping, with the intention that you subconsciously learn quantum mechanics. I have good reason to think this will be effective. You happen to have no interest in quantum mechanics and the facts that you learn have no bearing on your wellbeing.⁶⁵ Is this interference justified on balance? The argument in favour is that your knowing more about quantum mechanics is of greater value than respect for your personal sovereignty. It seems intuitive, however, that the loss to personal sovereignty is in fact a weightier concern than the gain in knowledge: indeed, it looks as though you would be morally correct to admonish me for my secretive interferences even if you wanted to learn about quantum mechanics. The strict epistemic paternalist needs to explain why this intuition is faulty.

Importantly, some forms of strict epistemic paternalism are less plausible than others. An ‘extreme’ strict epistemic paternalist holds that the promotion of epistemic value is the *only* good that can out-balance personal sovereignty (considerations of wellbeing never provide justifying reasons for paternalistically interfering). In order to defend this extreme view the epistemic paternalist must show not only why the promotion of epistemic value sometimes outweighs considerations of personal sovereignty, but why it is the only consideration that can. In other words, an explanation is needed as to why the value of knowing quantum mechanics outweighs the need to respect your personal sovereignty, whilst considerations of your happiness do not.⁶⁶

A more moderate position holds that the balancing condition can be met with either the promotion of wellbeing *or* epistemic value. The moderate strict epistemic paternalist is thus also a general paternalist. Recall, however, that general paternalistic interferences permit bringing about an epistemic decline (section 3). In order to rule out such interferences, and thereby remain distinctive from general paternalism, the strict epistemic paternalist needs to provide not only an account of how epistemic value can trump personal sovereignty, but also how it counterbalances considerations of wellbeing when the two come into conflict. If it seems unlikely that epistemic value outweighs personal sovereignty, then it is even less likely that it also counterbalances considerations of wellbeing. Again, there is pressure on the epistemic paternalist to explain why this intuition is faulty.

Perhaps a fairer interpretation of the moderate position is that we are only justified in paternalistically interfering with an individual when it promotes both epistemic value *and* wellbeing. In order to hold this position, however, an explanation needs to be given as to why either value taken alone is insufficient for justifying paternalistic interferences. One option is to provide an account of how wellbeing and epistemic value are necessarily tied together in

paternalistic justification, without reducing epistemic value to wellbeing value (as this would amount to eudaimonic epistemic paternalism). Note, however, that this would render any general paternalistic interferences that bring about an epistemic decline as unjustified. This defence of strict epistemic paternalism would have the radical consequence of eliminating the possibility of justified general paternalism, because wellbeing alone will be insufficient for justifying paternalistic interferences. The strict epistemic paternalist thus still needs to explain why epistemic value is weightier than the wellbeing concerns of the general paternalist.

5. Conclusion

In this paper I have argued that the epistemic paternalist is faced with a dilemma: either there is no moral distinction between epistemic paternalism and paternalism in general, or a distinctive project of justified epistemic paternalism is implausible. My argument hinged on the claim that the epistemic paternalist needs to provide an account of epistemic value in order to motivate their project. For this, they have two options: either the value of epistemic improvements are explained in terms of wellbeing (eudaimonic epistemic paternalism), or epistemic value is taken to be independently valuable (strict epistemic paternalism).

First, I argued that if the motivation for epistemic paternalism appeals to the value of wellbeing, the position collapses into general paternalism. Given this, the eudaimonic epistemic paternalist must concede that paternalistic interferences that worsen an individual's epistemic situation are justified. This is because there are plausible cases in which an individual is better off (in terms of her wellbeing) for being epistemically worse off. I further argued that the eudaimonic epistemic paternalist cannot rule out interferences that bring about an epistemic worsening as

unjustified in a non-arbitrary way. A morally distinctive project of epistemic paternalism is thereby ruled out.

Second, I argued that whilst the strict epistemic paternalist can avoid the collapse into general paternalism this comes at the cost of its plausibility. Against the most extreme form of strict epistemic paternalism I argued that it is implausible that epistemic value is the only thing that can outweigh personal sovereignty. I also suggested two more modest views. On the first, the strict epistemic paternalist holds that personal sovereignty can be outweighed by considerations of either epistemic value or wellbeing. Like eudaimonic epistemic paternalism, however, this position needs to address the possibility of general paternalistic interferences that bring about an epistemic decline. In order to hold on to epistemic paternalism as a morally distinctive project, an account needs to be provided as to why epistemic value can counterbalance considerations of wellbeing when the two come into conflict. The second modest approach was that the only interferences that count as justified are those that are simultaneously epistemically and generally paternalistic. In order to defend this position, however, the strict epistemic paternalist needs to provide an account of why either value alone is insufficient for justifying paternalistic interferences. If successful, this would eliminate the possibility of justified general paternalism.

The second horn of the dilemma is not a devastating objection to the project of epistemic amelioration, but it is important. What I have shown is that if the project of epistemic amelioration is to be convincing, then the strict epistemic paternalist needs to provide a detailed account of not only the nature of epistemic value but also of the justificatory weight it bears. If on the other hand the epistemic paternalist prefers to stick with the eudaimonic account of epistemic value then she will have to give up thinking that there is anything morally distinctive about her project.

Acknowledgments

I would like to thank two anonymous reviewers for their constructive comments on a previous draft of this paper. Many thanks also to my colleagues at Central European University for fruitful discussions on early ideas for this paper, and to those who participated in my lecture on this topic at Durham University.

NOTES

¹ See Gerald Dworkin (2010) ‘Paternalism’, in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* <<http://plato.stanford.edu/archives/sum2010/entries/paternalism/>> accessed 2 November 2013 and Emma C. Bullock, ‘A Normatively Neutral Definition of Paternalism’, *Philosophical Quarterly*, 65 (2015a): 1-21

² For instance, S.V. Shiffrin, ‘Paternalism, Unconscionability Doctrine, and Accommodation’, *Philosophy & Public Affairs*, 29 (2000): 205–50; Gerald Dworkin ‘Paternalism’, in R. Sartorius (ed.) *Paternalism*, (Minneapolis: University of Minnesota Press, 1983) pp. 19–34; D. Scoccia ‘In Defense of Hard Paternalism’, *Law and Philosophy*, 27 (2008): 351–81; B. New ‘Paternalism and Public Policy’, *Economics and Philosophy*, 15 (1999): 63–83; T. Pope, ‘Counting the Dragon’s Teeth and Claws: The Definition of Hard Paternalism’, *Georgia State University Law Review*, 20 (2004): 659–722; Sarah Conly *Against Autonomy: Justifying Coercive Paternalism* (Cambridge: Cambridge University Press, 2012); C. Coons and M. Weber *Paternalism: Theory and Practice* (Cambridge: Cambridge University Press, 2013).

³ For a central defense of epistemic paternalism see Kristoffer Ahlstrom-Vij *Epistemic Paternalism: A Defence*. (Basingstoke: Palgrave Macmillan, 2013). See also, Kalle Grill & Sven Ove Hansson, ‘Epistemic Paternalism in Public Health’. *Journal of Medical Ethics*, 31, 11 (2005):648-653; Alvin I. Goldman ‘Epistemic paternalism: Communication control in law and society’. *Journal of Philosophy*, 88, 3 (1991):113-131; D. Pritchard ‘Epistemic Paternalism and Epistemic Value’. *Philosophical Inquiries*. 1, 2 (2013): 9-37; and Jeroen de Ridder, ‘Is there epistemic justification for secrecy in science?’ *Episteme* 10, 2 (2013):101-116.

⁴ Goldman op. cit., and Ahlstrom-Vij op. cit.

⁵ Pritchard op. cit.

⁶ Ahlstrom-Vij, op. cit., p.10

⁷ Throughout this paper I focus only on the moral justification of epistemic paternalism. For discussion of its epistemic justification see Goldman, op. cit.

⁸ See Ahlstrom-Vij, op. cit., p. 51.

⁹ This is not the same as requiring that the outcome aimed at by the intervention is contrary to what the individual takes to be for her own good. Rather, paternalism involves a failure to treat the individual as an authority on whether or not she ought to be interfered with at all. See Daniel Groll, ‘Paternalism, Respect, and the Will’ *Ethics*, 122, 4 (2012):692-720., p. 696 and Bullock, 2015a, op. cit. An outlier is VanDeVeer who relies on some form of consent for

paternalism to be justified. See D. VanDeVeer, *Paternalistic Intervention: The Moral Bounds of Benevolence* (Princeton: Princeton University Press, 1986)

¹⁰ The following examples are illustrative: they are not intended as exclusive or equally justified examples of paternalistic interference.

¹¹ J. Gardner, 'Ethical issues in public health promotion' *South African Journal of Bioethics and Law*, 7, (2014): 30-33. p31.

¹² See for instance, D. E. Ost, 'The 'Right' not to know', *The Journal of Medicine and Philosophy*, 9 (1984): 201-312; P. Malpas, 'The right to remain in ignorance about genetic information--can such a right be defended in the name of autonomy?' *Journal of the New Zealand Medical Association*, 118, (2005): 71-78; R. Rhodes, 'Genetic links, family ties, and social bonds: rights and responsibilities in the face of genetic knowledge' *Journal of Medicine and Philosophy*, 23, (1998): 10-30.

¹³ Emma C. Bullock, 'Mandatory Disclosure and Medical Paternalism,' *Ethical Theory and Moral Practice*, (Online First: 2015b) DOI 10.1007/s10677-015-9632-2

¹⁴ Simone Chambers, 'Behind Closed Doors: Publicity, Secrecy, and the Quality of Deliberation.' *Journal of Political Philosophy*, 12, 4 (2004):389-410, p. 392 See also, Goldman, op. cit., pp. 115-18

¹⁵ Goldman, op. cit., pp. 121-22

¹⁶ Grill and Hansson, op. cit., p. 650. Discussions of 'information overload' similarly reflect the concern that individuals are simply unable to digest large amounts of information, therefore making them epistemically worse off. K. E. Himma, 'The concept of information overload: A preliminary step in understanding the nature of a harmful information-related condition.' *Ethics and Information Technology* 9, 4 (2007):259-272, p.260

¹⁷ With thanks to Simon Rippon for this example.

¹⁸ I. Stewart & J. Cohen *Figments of reality: The evolution of the curious mind*. (Cambridge: Cambridge University Press, 1997), p. 187

¹⁹ For a defense of the claim that doxastic voluntarism is impossible see, Robert Audi 'Self-evidence', *Philosophical Perspectives* 13, s13 (1999): 205-228. For the claim that direct doxastic coercion is impossible see Benjamin McMyler, 'Doxastic coercion,' *Philosophical Quarterly* 61, 244 (2011): 537-557.

²⁰ This example is adapted from the original in McMyler, op. cit., p550

²¹ Nudge paternalism need not be restricted to bringing about improvements in welfare and might sometimes qualify as a form of epistemic paternalism. For instance a policy implementing a nutrition labeling system on unhealthy foods might aim at nudging customers to making healthier diet choices, but also nudge them towards developing an understanding about foods that generally count as healthy or unhealthy (this perhaps being avoidable should the individual prefer not to read labels). My thanks to an anonymous reviewer for pointing this out.

²² Sunstein and Thaler op. cit., 157

²³ Goldman, op. cit.; Ahlstrom-Vij, op. cit.

²⁴ Pritchard, op. cit.,

²⁵ The names for these distinctions vary wildly between authors. Here I follow Pope, op. cit.; Compare, however, Dworkin op. cit; J. Feinberg *The Moral Limits of the Criminal Law – Volume Three: Harm to self* (Oxford: Oxford University Press, 1986) and J. Kleinig *Paternalism* (Manchester: Manchester University Press, 1983).

²⁶ Cf. Pope, op. cit. p687

²⁷ Cf. Feinberg op. cit. p8; Pope op. cit. p699

²⁸ A. Barton ‘How tobacco health warnings can foster autonomy,’ *Public Health Ethics*. 6, 2 (2013): 207-219, p.215

²⁹ L. Zagzebski, *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief* (USA: Oxford University Press, 2012).

³⁰ Ahlstrom-Vij, op. cit. pp 130-154

³¹ A. Hazlett *A luxury of the understanding: on the value of true belief* (Oxford: Oxford University Press, 2013) p.2
For example, L. Zagzebski “Epistemic Value and the Primacy of What We Care About,” *Philosophical Papers* 33, 3 (2004): 353–77

³² Hazlett op. cit., p28. For example: M. David “Truth as the Epistemic Goal” in *Knowledge, Truth and Duty: Essays on Epistemic Justification, Virtue and Responsibility*, ed. M. Steup (Oxford: Oxford University Press, 2001). E. Sosa “For the Love of Truth” in *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*, eds. A. Fairweather and L. Zagzebski (Oxford: Oxford University Press, 2000)

³³ These examples of arbitrary paternalisms are perhaps a little facetious. Certainly true belief and understanding have a far more important and recurring role in our wellbeing than dental health. (My thanks to an anonymous reviewer for highlighting this.) The point here, however, is not that the aims of epistemic paternalism are on a par with the aims of dental paternalism. Instead the claim is that it is absurd to insist that paternalism is only justified in order to bring about one particular wellbeing good given that the value of that good lies in its contribution to wellbeing more generally.

³⁴ See Hazlett op. cit., p8 and D. Pritchard & J. Turri “The Value of Knowledge”, *The Stanford Encyclopaedia of Philosophy* (Spring 2014 Edition) Edward N. Zalta (ed.), URL= <<http://plato.stanford.edu/archives/spr2014/entries/knowledge-value/>>

³⁵ R. Barnett ‘Bad Trip: Drug Prohibition and the Weakness of Public Policy’. Review: Duke, S. B and Gross, A. C., (1993) America’s Longest War: Rethinking Our Tragic Crusade against Drugs. *The Yale Law Journal*. 103, 8 (1994):2593-2630, p.2603.

³⁶ S. B. Duke & A. C. Gross, *America's longest war: Rethinking our tragic crusade against drugs* (New York: G. P. Putnam’s Sons, 1993) p. 255

³⁷ *Ibid.*, p.257

³⁸ The epistemic paternalist might make two responses here. First, they could argue that the policy is unjustified because individuals are not as vulnerable to making epistemic mistakes as the policy makers suppose. Alternatively, they might argue that the deception or manipulation is justified in this case because the individual will be spared making epistemic mistakes caused by their cognitive biases: the policy in fact leads to an overall epistemic improvement. Neither response, however, provides a conclusive objection to the more general point that the example illustrates: we can just alter the example by stipulating that (1) the general public will in fact make epistemic mistakes

about drug use should they be given correct information and (2) that the policy will not lead to an overall epistemic improvement (indeed, there is no reason to think that the policy will necessarily bring about an epistemic improvement). My thanks to an anonymous reviewer for identifying these possible objections to this example.

³⁹ Jennifer C Jackson, *Truth, Trust and Medicine* (London, Routledge, 2001): 98-99

⁴⁰ D. E. Moerman & W. B. Jonas, 'Deconstructing the placebo effect and finding the meaning response'. *Annals of Internal medicine*, 136, 6, (2002): 471-476, p. 473

⁴¹ Daniel Groll, 'What You Don't Know Can Help You: The Ethics of Placebo Treatment.' *Journal of Applied Philosophy*, 28, 2 (2011): 188-202, p.199.

⁴² Bullock, 2015b, op. cit.

⁴³ See L. Colloca and F.G. Miller "The nocebo effect and its relevance for clinical practice" *Psychosomatic medicine* 73, 7: 598-603

⁴⁴ See S. M. Miller and C. E. Mangan "Interacting effects of information and coping style in adapting to gynaecologic stress: should the doctor tell all?" *Journal of Personality and Social Psychology* 45, 1 (1983): 223-236); S. M. Flanagan "Adverse Outcomes after Colposcopy" *Doctoral Thesis* (University of Birmingham, 2012) M.H. Mishel "Uncertainty in acute illness" *Annual Review of Nursing Research* 15, 1 (1997): 57-80; F. Cohen and R. S. Lazarus "Active coping processes, coping dispositions, and recovery from surgery" *Psychosomatic Medicine* 35 (1973): 375-389

⁴⁵ See S.E. Taylor, J. S. Lerner, D. K. Sherman, et. al. "Are self-enhancing cognitions associated with healthy or unhealthy biological profiles?" *Journal of Personality and Social Psychology* 85, 4 (2003):605–15

⁴⁶ The epistemic paternalist might respond that the placebo and nocebo cases are unproblematic since they involve a local, rather than a global, deception. To defend this they would need to provide an explanation as to why the disvalue of epistemic deterioration depends on the degree to which it interferes with the individual's overall epistemic set up. My thanks to an anonymous reviewer for raising this point.

⁴⁷ See R.H. Thaler & C. R. Sunstein *Nudge* (London: Yale University Press, 2008); S. Conly *Against Autonomy* (New York: Cambridge University Press, 2013)

⁴⁸ Ahlstrom-Vij, op. cit. pp.15-50

⁴⁹ Hazlett, op. cit., p65

⁵⁰ Ryan T. McKay and Daniel C. Dennett, 'The evolution of misbelief', *Behavioural and Brain Sciences*, 32 (2009): pp. 493-561, p505

⁵¹ See Goldman, op. cit., p124 and Ahlstrom-Vij, op. cit., p139

⁵² See Pritchard, 2013, op. cit., p.18; and Goldman, op. cit., p127

⁵³ Ahlstrom-Vij, op. cit., pp.132-33.

⁵⁴ Pritchard, 2013, op. cit., p17

⁵⁵ Ahlstrom-Vij, op. cit., p.134

⁵⁶ Bullock, 2015b, op. cit., Adapted from R. Rhodes, 'Genetic links, family ties, and social bonds: rights and responsibilities in the face of genetic knowledge, *Journal of Medicine and Philosophy*, 23 (1998): 10-30 p. 20).

⁵⁷ Ahlstrom-Vij, op. cit., p134-5.

⁵⁸ Ibid. p.108. What I mean by ‘epistemic autonomy’ is sometimes referred to as ‘epistemic self-sufficiency.’ See for instance, J. V. Encabo, ‘Epistemic merit, autonomy, and testimony’, *Theoria*, 23, 1 (2008):45-56, p.45. This concept has also been referred to as ‘extreme epistemic egoism’ (Zagzebski 2012 op. cit., p252).

⁵⁹ E. Fricker, ‘Testimony and epistemic autonomy’. In Jennifer Lackey & Ernest Sosa (eds.), *The Epistemology of Testimony*. (Oxford: Oxford University Press, 2006) pp. 225--253.

⁶⁰ Ibid. and Zagzebski 2012, op. cit., p252.

⁶¹ Zagzebski 2012 op. cit., p260.

⁶² Ahlstrom-Vij, op. cit., p127

⁶³ Feinberg, op. cit. p56. See also R. R. Faden and T. L. Beauchamp *A History and Theory of Informed Consent* (Oxford: Oxford University Press, 1986), p19

⁶⁴ cf. D. Scoccia “Moral paternalism, virtue and autonomy” *Australasian Journal of Philosophy* 78,1 (2000): 53-71

⁶⁵ Perhaps epistemic paternalism is only justified when the epistemic improvement is consistent with the individual’s informed or rational epistemic goals, in which case foisting knowledge of quantum mechanics on a recipient who has no interest in the subject, even at a deep level, would be unjustified. As it stands, however, the case is consistent with the defining conditions of epistemic paternalism (see section 1). If the case is disagreeable to the epistemic paternalist the onus is on them to explain why the interference is unjustified.

⁶⁶ It is reasonable to assume that the epistemic paternalist would not defend this extreme position, especially given that examples of epistemic paternalism tend to be muddled together with non-epistemic improvements: i.e. the limits on advertising protect consumer health; interferences with jurors promotes justice; public health campaigns encourage healthy living habits.