

II—WHAT'S WRONG WITH PATERNALISM: AUTONOMY, BELIEF, AND ACTION

DAVID ENOCH

Several influential characterizations of paternalism or its distinctive wrongness emphasize a belief or judgement that it typically involves—namely, the judgement that the paternalized is likely to act irrationally, or some such. But it's not clear what about such a belief can be morally objectionable if it has the right epistemic credentials (if it is true, say, and is best supported by the evidence). In this paper, I elaborate on this point, placing it in the context of the relevant epistemological discussions. I explain how evidentialism is opposed to such thoughts; I show that possible ways of rejecting evidentialism (along lines analogous to those of pragmatic encroachment) won't work; and I sketch an account of the wrongness of paternalism that doesn't depend on any flaw in the belief about others' likely behaviour.

I

Introduction. Some people are very much into finding a *definition*, or at the very least a set of necessary and sufficient conditions, for paternalism.¹ Conditions are discussed, counterexamples offered, revisions put forward, and the game goes on. This will not be my game: 'paternalism' is to a large extent a term of art, and the search for a definition for a term of art is even more futile than it is for natural-language terms. Still, thinking about some of the suggested definitions out there can be productive in focusing our attention on a feature that seems to many to be central to the vague, murky family of normative phenomena that are often thought of as cases of paternalism. This is what I want to do in this paper.²

¹ For a recent survey—helpful and quite comprehensive, as far as I can see—and one that also includes some explicit methodological discussion, see *Dworkin* (2013). See also *Bullock* (2015), and the references there.

² *Groll* (2012, pp. 694, 707, 708) and *Tsai* (2014, pp. 81, 85) offer somewhat similar methodological remarks. They focus on what makes paradigmatic cases of

The kind of characteristic that I have in mind is one that ties the very nature of paternalism, or at least its distinctive wrong-making feature, to a kind of a *judgement*, or *belief*. The thought is that all cases of paternalism necessarily involve, or anyway most of them typically involve, a disturbing judgement about the paternalized's competence (or some such), and that the presence of such a judgement plays a crucial role in explaining what is necessarily, or at least always, or often, or *pro tanto*, wrong with paternalism. (Because I am not playing the definition game, I don't have to decide between these different disjuncts.)

Jonathan Quong, for instance, in his critical survey of several suggestions for definitions of paternalism (2011, pp. 73–80), gives examples of what he takes to be paternalistic behaviour that does not involve a violation of (independent) rights, or liberties, or even intrusion on the paternalized's autonomy. If, for instance, I refuse to give you £50 which you're asking for, because I think you will misuse the money, this refusal is paternalistic (says Quong), even though the money is *mine*, and so my refusal does not violate a liberty of yours or intrude on your autonomy. Compare this to a case in which I don't give you the money simply because I'd rather spend it myself—depending on the details, this may not be very nice, but surely it's not paternalistic, nor is it (*pro tanto*) wrong in the way that paradigmatic cases of paternalism are (*pro tanto*) wrong. What explains the difference? Quong thinks the difference comes down to the *judgement* that the paternalistic behaviour incorporates—the judgement that you will misuse the money, or that you aren't competent to make adequate decisions of this sort. Based on such examples, Quong suggests his *judgemental definition* of paternalism, the relevant part of which reads, '[The paternalizing] A's act is motivated by a *negative judgement* about [the paternalized] B's ability . . . to make the right decision or manage the particular situation in a way that will effectively advance B's welfare, good, happiness, needs, interests, or values' (Quong 2011, p. 80). And Quong is not alone in including some such judgement as a crucial part of the wrongness of paternalism.³

paternalism wrong, not on the word 'paternalism'. But even this is too strong for my purposes—we don't want to rule out from the start the possibility that the word 'paternalism' fails to designate a normative kind; that is, it's possible that different cases of paternalism are *pro tanto* wrong for different reasons, and that some of them just aren't.

³ For similar conditions—in the context of different understandings of paternalism—see Shiffrin (2000, p. 218), Tsai (2014, pp. 86–7), Groll (2012, p. 718).

The thought, then, that something in the vicinity of what's characteristic of paternalism has to do with the *judgement* of the paternalizing seems very common, and as Quong's example shows, it has considerable intuitive appeal. But it should also give you pause. For paternalism, if it is wrong, is *morally* wrong; but it is far from obvious that beliefs or judgements can be morally wrong at all. The norms that (directly) govern beliefs are *epistemic* norms, not *moral* (or other practical) ones. And so, if you want to include something about the relevant judgements in your characterization of paternalism or what is typically objectionable about it, you should have something to say about the relations between epistemic and moral norms, or between the moral status of beliefs and action. Such discussions are now quite common in the literature, but not, as far as I am able to determine, in the literature on paternalism. This is where this paper comes in.

My conclusions are tentatively negative: while such negative judgements about the agential abilities of the paternalized may be involved, their role in explaining the *pro tanto* wrongness of paternalism is different from the one suggested by the above quotes, and at the end of the day is fully explained by practical, not epistemic, norms.

My discussion proceeds as follows. In §II, I explain the objection to the centrality of the belief about others' likely behaviour in an account of the wrongness of paternalism, tying it to epistemological discussions of evidentialism. The following two sections discuss two related ways of rejecting evidentialism (and the said objection with them)—first, by invoking moral norms that directly govern the relevant beliefs (in §III), and then, in §IV, by allowing moral considerations to influence the relevant epistemic standards that govern the beliefs, in the way that is discussed in the context of so-called pragmatic encroachment. By the end of §IV, then, the evidentialist objection from §II will have been vindicated, at least *vis-à-vis* these two ways of rejecting it. A problem remains, however: the paternalized will often care most about the paternalizing's *belief*, and will resent it at least as much as the relevant paternalistic *action*. This creates a gap between what, on my suggestion, grounds the wrongness of paternalism and what the person wronged cares most deeply about. I discuss this problem in §V. In §VI, by way of conclusion, I sketch how we can account for the *pro tanto* wrongness of paternalism without assigning the beliefs about others' likely behaviour the kind

of role that the discussion in this paper rules out—the sketched account relies, unsurprisingly, on the value of personal autonomy, and shows (using the idea of an exclusionary reason) how this value can explain why you are sometimes not justified in acting on your negative belief about others' future behaviour—even when there's nothing objectionable about that belief itself.

We need one more preliminary. A distinction is sometimes drawn between interpersonal paternalism (as in Quong's example above) and institutional or political paternalism (when the paternalizing agent is the state, or the law, or some such). And we should keep an open mind about the possibility that there are important moral differences between the two. As we shall see, this distinction is indeed relevant to some of the points made below. For the most part, though, I intend to discuss both interpersonal and institutional paternalism. I explicitly draw on the distinction between them where this is relevant.

II

The Simple Evidentialist Objection. Suppose the available evidence indicates rather clearly that I'm the tallest member of my department. Suppose, for simplicity, that this is indeed so. What, then, should I believe regarding who the tallest member of my department is? The relevant 'should' is naturally read as an epistemic 'should', the kind that is fully determined by the evidence regarding relative tallness, or the truth of the target proposition, or some such. Notice also that this 'should' is essentially impersonal—so long as the relevant evidence is available to you and me alike, both of us should have the same belief here—and that morality seems entirely beside the point.

Now suppose that the evidence indicates rather clearly that I'm the smartest member of my department, and suppose also, for simplicity, that this is indeed so. What, then, should I believe regarding who the smartest member of the department is? True, things are more complicated here, for reasons we will get to shortly. But isn't there an obvious appeal to the thought that here too, the relevant 'should' is naturally read as an epistemic one, and so I should believe the proposition that's both true and most supported by the evidence, namely, that I'm the smartest person in my department? Clearly, if

this is both true and best supported by the evidence, this is the thing *for you* to believe, isn't it? And if so, why isn't it the thing *for me* to believe as well, if the same evidence is available to us? And if the question is epistemic, shouldn't morality here too be beside the point?

Now, it seems that it isn't. Believing that I'm the smartest member of the department—unlike that I'm tallest—seems *arrogant*, and so morally objectionable, even when best supported by the evidence.⁴ And while it seems objectionable for me to so believe, it does not seem objectionable for you to believe that I'm the smartest member of the department, even if we are privy to exactly the same evidence.

But perhaps we can accommodate these phenomena in non-radical ways—without rejecting, that is, the thought that we should believe what is best supported by our evidence.⁵ Perhaps, for instance, what is (sometimes) morally objectionable is not exactly the belief that one is the smartest around, but rather thinking about this question to begin with. Perhaps, in other words, there is no moral flaw in coming to believe one is smartest, *if one already asks oneself whether this is so* (or some related question), and if this is the conclusion that the evidence supports; it's just that it's non-virtuous to ask such questions to begin with. Or perhaps I should safeguard myself against self-serving biases: knowing one or two things about human psychology, I should know that I am likely to interpret the evidence in a self-serving way, and so I should endorse some epistemic policies that deal with such tendencies—say, if the evidence seems to me to support an arrogant conclusion, perhaps I should discount it somewhat, and reduce confidence in the arrogant conclusion. And interestingly, the 'shoulds' in the last few sentences may be moral, or epistemic, or both. And certainly, moral norms (for instance, about the objectionableness of arrogance) may dictate that I shouldn't *say* that I'm the smartest around, perhaps that I shouldn't act on the (true) belief that I am, perhaps that I should *act as if* I am not.

The question I want to focus attention on now is distinct from these other ones, and is in a sense cleaner. Having taken into account biases and attempts to compensate for them, and given that I am going to form a belief regarding my comparative intelligence, and that

⁴ For a discussion of modesty that accepts believing (to an extent) against the evidence as morally called for, see Driver (1999), and the references there.

⁵ This, I think, is a good general characterization of Kawall's (2013) response to Stroud (2006) and Keller (2004, 2007).

if it turns out I'm the smartest around I'm not going to say so, or act on this belief, indeed I will go to great efforts to forget that this is so—given all this, if the evidence strongly indicates that I'm the smartest person around, isn't that what I should believe? Isn't this cleaner question entirely epistemological? And if it is, isn't morality silent about it?

Evidentialists answer in the affirmative. Evidentialism, roughly, is the view that whether one should believe a proposition, or at least whether one *epistemically* should believe it, is entirely determined by the evidence.⁶ In evidentialist terms, then, there's no difference between the two cases I started with—just as I should believe that I'm the tallest person around when this is what the evidence (sufficiently strongly) indicates, I should believe that I'm the smartest person around when *this* is what the evidence (sufficiently strongly) indicates. Evidentialists need not be blind to the intuitive difference between the two cases, but they will claim that these differences can be fully accommodated by the kind of explanations offered in the paragraph before last. When the cleaner question is clear, its answer is fully determined by the evidence. So, at least, say evidentialists, and it can't be denied that their answer is extremely plausible.

Back to paternalism, then. You are considering your friend's request to borrow £50. What should you believe about her competence to spend it well? The natural, evidentialist answer is as straightforward as it is non-moral: what you should believe is what is best supported by the evidence. If the evidence indicates that she will likely make poor decisions, or fail to follow through on her good ones, or some such—then this is what you should believe. You are not wronging her, or failing to treat her with respect, or failing to respond appropriately to her Kantianly rational nature, by believing about her what you should believe about her, namely, what the evidence sufficiently strongly indicates. Of course, perhaps you shouldn't believe such things of her. Perhaps, in other words, the evidence doesn't support these beliefs, perhaps it even supports more

⁶ See, for instance, Mittag (2004), and Conee and Feldman (2004). Conee and Feldman's characterization, unlike Mittag's, includes the 'epistemically' qualifier. So Conee and Feldman's formulation—unlike Mittag's—leaves it open that whether, say, one is *pragmatically* or even *morally* justified in believing a proposition may depend on factors other than the evidence. One way of understanding this kind of option—as about terribly important consequences that will follow if one does (or does not) believe something—makes it irrelevant for our purposes. Other ways make this option very relevant—as will be discussed in the next two sections.

optimistic beliefs. If so, your negative judgements about your friend's likely behaviour are epistemically flawed, and again morality is irrelevant. But the thought that there is a moral flaw in forming these negative judgements about her when they are best supported by the evidence lies somewhere between falsehood and a category mistake.⁷

Similarly for other cases of paternalism. Whether the state should nudge people into saving more for retirement is a complicated moral and political question, but whether we should *believe* that most of us are highly likely to under-save is an epistemic question, and the evidence seems to show clearly that the answer is 'yes'.⁸ Whether the state should make alcohol unavailable to the alcoholic is a complicated moral and political question,⁹ but what we should believe about the likely results of alternative policies here is an epistemic question, fully answered by the evidence, not by moral considerations about the alcoholic's dignity. Whether my wife should hide the candy from me is an interesting moral question, but what I am likely to do if she does not is a non-moral factual matter, and what she should believe about it is determined by the evidence. She may wrong me by hiding the candy—but she does not wrong me by forming the well-supported belief that if she doesn't, I will eat much too much of it.¹⁰

Of course, here too we can tell other, related stories. Perhaps, for instance, while my wife doesn't wrong me in believing according to the evidence here, she would have wronged me had she talked about this to my students. Or perhaps state officials suffer from some interesting biases—perhaps belittling the wisdom and strength of will of the common man and woman—and should (epistemically as well as morally) do what they can to compensate for such biases, so as not to believe too easily that we are unlikely to save sufficiently for our

⁷ At one point, Quong (2011, p. 102) anticipates something close to this objection. But I don't see how what he has to say in response is at all relevant to the problem in the text here, and this raises the suspicion that the problem he has in mind is not the one I elaborate on here.

⁸ See Thaler and Sunstein (2008, pt. 2), and the references there.

⁹ See, for instance, Mill (1859, pp. 167–70).

¹⁰ There is some literature on the relations between *trusting* and epistemic rationality (see, for instance, McGeer 2008, and the references there). Some of it may be relevant here—you may think, for instance, that my wife should *trust* me regarding the candy (or, more plausibly, regarding more important stuff). But I don't think going for trust here will save the thought that the problem with paternalism is the belief—for one thing, some of the worries I get to below can be restated in terms of trust as well. I hope to discuss trust and related phenomena in future work.

retirement. Perhaps of special importance in the political sphere is when the relevant negative judgement is directed at a ‘suspect category’—we may very well be especially biased in our epistemic tendencies to believe, say, that people of certain minorities are weak-willed, or perhaps we tend to believe that a woman is likely to be irrational on the kind of evidence on which we would not have concluded that a man was likely to be irrational, and perhaps, in general, we are less likely to listen to and believe minorities, or the less well educated, or the disempowered. If so, these phenomena clearly have epistemic and moral significance, for instance, in the kind of measures we are required to undertake in order to fight such biases.¹¹

Still, when all of this is said and done, and when we’re focusing on the cleaner question about whether, say, my wife should believe that I am likely to be weak-willed (again) and eat too much candy, or whether we should all believe that we are likely to under-save for retirement, these questions are fully answered by the evidence. And when someone forms a negative judgement about you (that you will under-save, that you will eat too much candy) that is most strongly supported by the evidence, they believe as they should, and they do not wrong you in any way. Paternalism may yet be (*pro tanto*, at least typically) morally wrong, of course. But its wrongness cannot be grounded in this direct way in the moral wrongness of the relevant judgement or belief.

So, at any rate, says the evidentialist. Before proceeding to see whether evidentialism can be qualified in a way that will challenge this conclusion, I want to note another, related point.

Among judgements about the potential paternalized, we should distinguish—as some writers sometimes fail to do—between negative judgements about their relevant *abilities* or *competences*, and judgements about their *likely behaviour*. In his definition (quoted above), for instance, Quong speaks of ‘a negative judgement about [the paternalized] *B*’s *ability* . . . to make the right decision or manage the particular situation in a way that will effectively advance *B*’s welfare, good, happiness, needs, interests, or values’ (2011, p. 80). But I think this is a mistake,¹² and probably not a deep one—it seems to me that even Quong would agree to classify as paternalistic, and as *pro tanto* objectionable for roughly the reasons paternalistic actions

¹¹ Such phenomena—and their implications—are widely discussed in Fricker (2007).

¹² I find a similar mistake in Tsai (2014).

are *pro tanto* objectionable, an intervention that satisfies his other conditions, and that is motivated by a negative judgement, not about *B*'s *ability*, but rather about *B*'s *likely behaviour*. This is different because, obviously, *A* can consistently think of *B* both that *B can* exercise good judgement and strong will, and that he is *unlikely* to. And this difference may make a difference because questioning someone's *ability* to deliberate and act rationally seems much more radical an insult to his or her nature as a rational agent, than merely questioning—on empirical grounds—his likelihood of making, in a specific setting, the right decision, and then following through. This becomes especially important when thoughts about one's abilities and competence as a Kantianly rational agent are naturally related to thoughts about *moral status*, so a negative judgement about one's abilities and competences may amount to a failure to fully recognize one's moral status (as a person, say). But when, having reviewed the empirical evidence about how we tend to discount our future interests, I conclude that you (like me) are likely to under-save for your retirement, I am not in any way failing to acknowledge your moral status as a person.¹³ And as this last example shows, there needn't be anything condescending about it either (for I fully acknowledge that I too am subject to the same likely flaws).¹⁴

Now, even when it comes to beliefs about people's abilities, the previous points seem to apply—whether or not someone has the relevant abilities is a factual question, and whether or not we should believe this is, arguably, fully determined by the evidence. Above I spoke of questioning someone's rational abilities and competences as failing to *recognize* their nature as persons, but the use of the word 'recognize' amounts to cheating here. You can only recognize what's there, and if the evidence shows that someone lacks the abilities constitutive of personhood, this, arguably, is the thing to believe about him or her. When it comes to the distinction between epistemic and moral norms, beliefs about people's abilities are not significantly different from beliefs about their likely behaviour. Still, differences remain. Perhaps, for instance, when it comes to people's abilities, we should treat them *as if* no evidence supports the

¹³ For a similar point—made in the context of criticizing Waldron's (1999) claim that the distrust of the common folk that is implicit in the practice of judicial review is inconsistent with their moral status as acknowledged by the very rights judicial review is supposed to defend—see Enoch (2006, pp. 27–8).

¹⁴ See de Marneffe (2006, p. 79) for a related point.

judgement that these abilities are seriously compromised. I am not sure how plausible this thought is (and I return to related thoughts in the final section), but anyway, this thought is much more plausible when it comes to beliefs that are closely related to moral status than when the beliefs involved are empirical beliefs about the likely behaviour of certain people in certain circumstances. And because for our purposes here—the discussion of paternalism—talk of people’s abilities is just a red herring, from now on I will focus on the case of negative judgements about likely behaviour.

III

Can Rejecting Evidentialism Help? First Attempt: Moral Norms Directly Governing Beliefs. But you may work in the opposite direction. You may want to hold on to the thought that I should *not* believe that I’m the smartest person in my department, and perhaps also to the thought that when forming negative judgements or beliefs about how people are likely to behave, we do wrong them. And noticing that this is inconsistent with plausible versions of evidentialism, you may think of this as a reason to reject evidentialism.

The question, of course, is whether this can be done in a plausible and relevant way. And because evidentialism is so plausible, if you want to reject it you had better offer quite compelling reasons—reasons that accommodate its appeal, perhaps explain it away, offer a plausible substitute, and so on. And you are going to want to do all that in a way that is not too ad hoc—that is, not merely relying on the force of the purported intuitions about the nature of paternalism and its wrongness. I can think of two attempts here, and they are the topic of this and the next section.

On the first suggestion, moral norms may sometimes govern beliefs directly. As my point of departure here I’m going to take Sarah Stroud’s thoughts about friendship and epistemic partiality (Stroud 2006).¹⁵ Stroud notices how our epistemic responses to, say, disturbing evidence about a friend’s behaviour differ—and more importantly, *should* differ—from our responses to similar evidence about a stranger. The ‘should’ here is important, because on reflection, we

¹⁵ Keller (2004) makes very close points. It’s not obvious how to understand Stroud. She may be understood along the lines of the next section as well.

don't see this difference in response as a failure of rationality (like some kind of weakness of will, perhaps). We see it as morally desirable, virtuous, perhaps obligatory, a constitutive part of the very practice and relationship of friendship. Thus, we will tend to look for more exonerating evidence in the case of the friend; we will tend to require more incriminating evidence before we are willing to commit to a negative conclusion; we will tend to interpret evidence in a more positive way; we will invest intellectual (and sometimes other) resources in coming up with explanations of the data that are less disturbing; and at the end of the day, we will tend not to go in for negative judgements about our friends, even when the evidence available to us would have made a similar judgement quite reasonable in the case of a stranger. Indeed, within friendships, we seem to be *entitled* to such attitudes from our friends, and we would see it as a failure in a friend (or a failure of the friendship) if she were willing to treat the damning evidence against us without this kind of partiality.

Perhaps, then, something similar can be said about the case of judgements that are (arguably) constitutive of paternalism and its *pro tanto* wrongness. Perhaps just as moral norms of friendship govern directly what we should believe about our friends (given a specific body of evidence), moral or political norms of some other kind govern beliefs about people's likelihood of deliberating poorly, or of exemplifying weakness of will, and the like.

Stroud's suggestion amounts, of course, to a rejection of evidentialism, at least in full generality—when it comes to friendship, and perhaps in other cases too, perhaps including the paternalism one, what one should believe is not fully determined by the evidence. It is also determined by the nature of the relevant relationship and the moral norms governing it. And it's the kind of rejection of evidentialism we were looking for—it is motivated independently of the paternalism case, based on intuitively plausible judgements about fairly central, unrelated cases.

Still, I don't think this line can save the thought that what is characteristic of (the wrong-making feature of) paternalism is the judgement about another's likely failed decisions or behaviour, for the following two reasons.

First, Stroud's discussion sometimes fails to distinguish sufficiently clearly between different questions. Thus, the question of how much to invest in looking for exonerating evidence or for more charitable interpretations of the already available evidence, while of course

epistemically significant, is not cleanly *epistemic* at all. Perhaps we have a moral obligation to *look for* more exonerating evidence (looking for evidence being an action, not a belief), but what we should *believe*, at any given moment, is a function merely of the evidence available to us at that moment. The moral obligation to look for more evidence, then, is strictly speaking beside the point (or anyway, beside our point here). In this respect, many of the phenomena Stroud discusses are analogues of the non-clean, non-radical cases mentioned in the previous section. Perhaps, returning to the paternalism case, our duties towards our fellow citizens require that we invest resources in trying to de-bias ourselves (and our political system), to reduce the effect of prejudice and self-serving biases, to make sure—before we paternalize someone—that they really are unlikely to make good decisions, and so on. These are all practical matters, not epistemic ones (though they may be epistemically significant, of course). The cleaner question is whether, given a body of evidence, what we should believe (whether our fellow citizens are likely to under-save for their retirement, say) is influenced by anything other than the evidence.

Stroud expresses intuitions about this cleaner question as well. But I don't think she ever formulates the relevant kind of case in sufficiently precise details. The kind of case I have in mind is one where all the relevant evidence about a person is in; I then justifiably form a negative judgement about his behaviour in some affair; I then find out that the person whose behaviour we've been discussing is my friend; and I then proceed to withdraw the relevant judgement. (Or similarly in the opposite direction: the evidence about my friend is in; I refuse to form the negative judgement; I then find out it's really about a stranger; and proceed to go in for the negative judgement.) And my first objection to the attempt to use Stroud in our context is that once the cleaner case is described, Stroud's claims become much less plausible. It does not seem that, *all other things being equal*, we should respond differently in our beliefs in the case of the friend and the case of the stranger. At the very least, if the friendship gives us a reason to respond differently in this case, then it's the wrong kind of reason, similar to the reason we would have to believe something from an offer to pay us if we did.¹⁶ If the clean paternalism case is

¹⁶ At one point Stroud notes the relation to the 'wrong kind of reasons' problem (2006, p. 513 n. 30). She doesn't discuss it at length, nor does she notice that it's especially troubling in the kind of case discussed in the text.

supposed to be modelled on the friendship one in this way, then the argument may fail at its starting point.

Let's assume, though, for the sake of argument, that with regard to the 'clean' friendship case too, Stroud is right. Let's assume, in other words, that holding everything else constant, a belief may be justified in the stranger case that is not justified in the friend case. My second objection to utilizing the Stroud line in the paternalism case is that even if it is plausible in the friendship case, it is much less plausible in the paternalism case—at least in the case of political, institutional paternalism.

The friendship case is a case of *partiality*, as Stroud emphasizes (even in her title). It is grounded in the nature and value of a special, close, and non-universal relationship—that between you and specific others, others who are special to you. Whatever plausibility thoughts of the epistemic relevance of the moral norms have here it owes to these features of the friendship case. But these features are not shared by the case of political paternalism. There, the moral norms that are supposed to govern the belief (in the projected irrationality or akrasia of some others, say) are not partial, they are universal, and to call the relation between one and one's fellow citizens a close relationship would be a huge stretch (and a dangerous one too).¹⁷ Perhaps, in other words, there is some plausibility to the thought that 'Friendship requires epistemic irrationality' (Stroud 2006, p. 518). The thought that *politics* requires epistemic irrationality is almost beyond belief.

How about interpersonal paternalism? The thought that among my wife's relationship-based obligations to me there's also an obligation not to form negative judgements about me even in the face of evidence that would have sufficed in the case of a stranger is not as implausible as is the political case—it's just the friendship case again.¹⁸ In this respect, then, things look better for the relevance of the judgement or the belief in interpersonal paternalism than they do for political liberalism. However, first, one may hope for a unified account of (the wrong-making feature of) paternalism, and so the

¹⁷ It is perhaps less of a stretch to believe that *some* partiality is owed to one's fellow citizens, that, say, one owes more to one's fellow citizens than to others. This too seems beside the point, though—the inappropriateness of the judgements purportedly constitutive of paternalism does not seem to depend on whether the paternalized is a member of one's own political community.

¹⁸ It's a different close relationship, of course, but I don't think the difference makes a difference in our context.

problems in this way of accounting for political paternalism project onto the case of interpersonal paternalism as well. And second, now focusing on just interpersonal paternalism, it's important to note that the moral norms that govern intimacy may work in both directions. Thus, perhaps my wife is under special obligations not to think ill of me. But she is also under special obligations, compared to strangers, to promote my wellbeing, and the privacy constraints that apply to her interfering in my affairs are much more lax than they are for strangers. This is why, whatever you may think about her hiding the candy (in order to promote my wellbeing and protect me from my own akrasia), surely you think it would be much more morally objectionable for a distant colleague of mine to do so. So it would be an oversimplification—and a tendentious one—to think that the intimacy of the interpersonal relationship only explains the *pro tanto* wrongness of the interpersonal paternalistic intervention—depending on circumstances, it may actually explain its permissibility.

Much of what Stroud has to say about epistemic partiality, then, is both deep and intuitive, and it may require revisions in more traditional ways of thinking about epistemic rationality. But none of this will help the attempt to ground the *pro tanto* wrongness of paternalism in the negative judgement about the paternalized's likely behaviour.

IV

Can Rejecting Evidentialism Help? Second Attempt: Moral (and Political) Encroachment. But perhaps the relation between the moral considerations and the epistemic ones is somewhat more subtle than that. Perhaps epistemic standards are somewhat flexible, and perhaps pragmatic—including moral—considerations can determine, within the general epistemic flexibility, the relevant more precise epistemic standards.

Such is the underlying thought of what is sometimes called 'pragmatic encroachment'. According to it, pragmatic considerations, such as what is at stake, can (constitutively, directly, not causally) make a difference to whether someone knows something.¹⁹ Such

¹⁹ I borrow this characterization of pragmatic encroachment from my 'Political Philosophy and Epistemology' (Enoch forthcoming).

views are often motivated by examples in which we are comfortable with knowledge attributions when the pragmatic stakes are low, but not so when the stakes have been raised, and this even though evidence has been held fixed throughout.²⁰ One common example is that of so-called bank cases, where, based on the same evidence (remembering going to the bank and making a deposit on a Saturday, a couple of weeks ago), one is said to know that the bank will be open tomorrow, on Saturday, when the stakes are low, but once the stakes are high (if a deposit is not made by tomorrow, we will lose our house), we tend to retract the knowledge attribution. Perhaps, then, the strength of evidence needed for a true belief to amount to knowledge varies with context, and in particular with some pragmatic features of the context, as in bank cases. Thus, one way of understanding the thought underlying pragmatic encroachment is that there's a *very* close connection between knowledge, or what we are justified in believing, and *action*.²¹ And because what it is permissible to take as a premiss in our practical reasoning varies with contextual pragmatic features, so does the relevant standard for (epistemic) justification or knowledge.

A natural way of extending thoughts of pragmatic encroachment is to think about how what is *morally* at stake can affect the relevant epistemic standards, or, in other words, about *moral encroachment*.²² Perhaps, for instance, the question how strong the evidence must be if the belief is to be outright (epistemically) justified (and, if all goes well, even amount to knowledge) gets a different answer not just in contexts that differ with regard to how high the *prudential* stakes but also with regard to how high the *moral* stakes are. Notice that, on this thought, moral norms do not directly govern beliefs. Beliefs are directly governed by epistemic norms alone. It's just that

²⁰ See Hawthorne (2004), Stanley (2005), and Ichikawa and Steup (2012, §11), and the references there.

²¹ 'The basic idea of advocates of pragmatic encroachment is that one is epistemically justified in believing a proposition only if one is pragmatically justified in acting as if *p* is true (using *p* as a premise in one's practical reasoning)' (Pace 2011, p. 256).

²² I take the term from Pace (2011), on whom I rely extensively in this section. However, I abstract the discussion from some of his context and commitments: Pace's discussion is conducted in the context of interpreting, and to an extent vindicating, James on theism and the will to believe, but none of that is relevant here. Also, while Pace rejects evidentialism about justified beliefs, he accepts evidentialism about justified credences or degrees of confidence (2011, p. 259). For reasons I can't go into here, I find this combination problematic. The moral encroachment thought I present in the text is not committed to this combination of views.

the content of the epistemic norms (or some of them) is somewhat sensitive to moral (and other practical) concerns.

Think about some of our examples again. Perhaps, for instance, the thing to say about forming a negative judgement about a friend's behaviour is that this should be based on the evidence, as in the case of a stranger; it's just that the moral stakes—those having to do with the value of friendship—affect the threshold of evidential support that justifies outright belief and that is needed for knowledge. If so, based on the same body of evidence regarding the friend and the stranger, it's possible that, say, refusing to form the negative judgement is criticizable in the case of the stranger (because there's evidential support above the threshold relevant in *that* moral context) but not in the case of the friend (because the morally rich context of friendship raises the threshold of evidential support needed for a justified belief). Or think about the arrogance case again—perhaps the evidence needed for *me* to justifiably believe that I'm the smartest member of my department is stronger than the evidence needed for *you* to justifiably form that belief (namely, the belief that I'm the smartest member of our department). If so, it's still possible that I too should believe this—if the evidence is overwhelmingly strong. But, of course, in real life it hardly ever is. And anyway, moral encroachment can explain the initially puzzling asymmetry between the justification of my belief, that I'm the smartest member of my department, and yours.

In many contexts, then, thoughts about moral encroachment look very promising in accommodating intuitive judgements about the relation between beliefs and moral norms. This is even true in politics: in the context of defending a kind of a public-reason account of political legitimacy, for instance, Tom Nagel (1987) talks of a politically motivated higher epistemic standard needed in order to know, or justifiably believe, a principle that is going to serve as grounds for coercive political action. One natural way of understanding such talk is as talk of moral, indeed political, encroachment, in just the sense we've been discussing.²³

Perhaps, then, this is a promising way of understanding the case of paternalism as well. Here is how it would go. Our beliefs about

²³ For elaboration, and ultimately rejection, of this line of thought, see Enoch (forthcoming). It's perhaps worth mentioning that Nagel's paper was written about fifteen years before talk of pragmatic encroachment became more common in epistemology, so he can't be faulted for not addressing this literature.

the likely behaviour of others (and indeed, of our future selves as well) should be based on the evidence, of course, not on their moral status or some such. But the threshold of evidential support needed for some relevantly important epistemic standard—justification, say, or warrant, or knowledge, or some such—does indeed depend on their moral status. So, perhaps because people are Kantian rational agents, or are owed a special kind of respect, the level of evidential support needed in order to justifiably believe that they will make a poor decision, or will be affected by weakness of will, is much higher—and perhaps even different in kind—than the evidential support that would suffice for justification in other cases, cases where the moral stakes, as it were, are lower. And perhaps this is especially so in the case of political paternalism. Perhaps, in other words, in the political context we are *especially* morally required to think of each other as, say, fellow inhabitants of a possible Kingdom of Ends; or perhaps at least the state is required to think of its citizens in this way. And while this is not *evidence* that people will make the right decisions and then follow through on them, this does raise the threshold of evidential support needed to justifiably believe that they will not.

There is a lot to like in this way of thinking about moral and political encroachment—and I suspect that this is the best way of fleshing out intuitions about the relevance—to the objectionability of paternalism, for instance—of beliefs and their epistemic status. And I think it needs to get more attention in political philosophy. Still, I don't think that it can save the thought that the distinctive *pro tanto* wrongness of paternalism—much less the very nature of paternalism—is ultimately epistemic, for the following two reasons.

First, there seems to be a mismatch here between the scope of the *explanandum* and that of the *explanans*. Even if it's true that moral (and other practical) stakes affect the location of the justification threshold (for instance), there are presumably purely epistemic bounds on such encroachment. Pace (2011, pp. 245–6), for instance, thinks that while pragmatic and moral considerations may make epistemically permissible a belief that otherwise wouldn't have been, they never do so if the belief is less supported by the evidence than its negation. Similarly, Stroud (2006, throughout) emphasizes that while friendship requires some epistemic partiality, it most certainly does not require epistemic blindness, and Driver (e.g. 1999, p. 827) emphasizes that while modesty may require underestimation of the

evidence about one's own advantages, it does this to a limited extent, and is not to be equated with self-deprecation. Perhaps, then, considerations of moral encroachment and how we ought to think of each other explain why we need *more* evidence to justifiably believe that people will under-save for retirement than we need for other beliefs, and why my wife needs stronger evidence to justifiably believe that I will succumb to candy temptation than she would for other beliefs. The striking thing, though, about many cases of paternalism is that the evidence about the likely bad decision or akratic action (or some such) is overwhelmingly strong, as indeed it is in the two cases I just used as examples. Even if moral considerations raise the threshold of epistemic support needed for justification, then, in many cases of paternalism we are easily over the raised threshold as well. My wife justifiably believes—indeed, she knows—that I will eat too much of the candy; and we all justifiably believe—indeed, we know—that without some intervention the vast majority of us will fail to save sufficiently for our retirement.²⁴ If so, at least in many paradigmatic cases, nothing about moral encroachment shows the judgements supposedly constitutive of paternalism in a bad light.²⁵

Second, and more importantly still, we need to keep our eye on the dialectical ball. The initial thought we're trying to pinpoint and develop is that something about *judgements*, perhaps at the end of the day something *epistemic*, plays a crucial role in understanding the nature of paternalism, and anyway, of the *pro tanto* wrong associated with it. In our attempt to make good on that thought, we have now been going for moral encroachment. But moral encroachment explains the problematic nature of some relevant belief—that we will under-save for retirement, that I will eat too much candy—not ultimately by doing epistemology, but by emphasizing practical, indeed moral, concerns (perhaps concerns about the value of autonomy, as I'm about to suggest). We seem to have come full circle,

²⁴ This raises a whole other issue—whether it is epistemically, and morally, and even legally acceptable to base judgements about a specific agent based on such statistical evidence. For some discussion, see Enoch, Spectre and Fisher (2012), and Enoch and Fisher (2015), and the references there.

²⁵ Gregg Strauss and Scott Shapiro (independently) suggested to me that perhaps we should distinguish between cases of paternalism in which the *pro tanto* wrongness is grounded in the relevant belief failing to reach the (partly morally determined) threshold needed for justification and knowledge, and cases in which we are above that threshold, and the *pro tanto* wrongness is grounded in the kind of practical considerations I mention in the final section below. I am not sure, but there may be interesting implications of the distinction between these two kinds of paternalism.

then: if we want to understand the *pro tanto* wrongness of paternalism, we need, on the current suggestion, to focus on the moral considerations that encroach on the epistemic, we need to understand the moral considerations *in virtue of which* the threshold of evidential support is higher in those cases than in others. And if so, it seems that talk of the relevant judgements and their inappropriateness was just a detour—at the end of the day what explains the wrongness of paternalism is not the poor judgements about the paternalized's likely behaviour or the inappropriateness, in some sense, of such judgements, but rather the moral considerations that render them inappropriate. You can still talk of the judgements if you want, but there will be something misleading about so doing, and focus on them will be at the very least less than fully perspicacious, for according to moral encroachment theory as well, what really does the work is the underlying moral, entirely practical, considerations (perhaps of the kind discussed, in a preliminary way, in the final section below).

Pulling the threads of this and the previous sections together, a dilemma emerges: impressed with intuitions about the epistemic relevance of moral considerations, you can either go in for a view according to which the moral norms govern beliefs directly, or for a more nuanced view of moral encroachment. According to the former, the explanatory role of the beliefs (about others' behaviour) and their moral status is indeed central to understanding the *pro tanto* wrongness of paternalism, but the thought that the moral norms govern directly the relevant beliefs seems implausible, especially in our context. Thoughts about moral encroachment seem much more subtle and plausible. But according to them, the beliefs pretty much drop out of the explanatory picture. Either way, then, there's trouble for a view that locates what's distinctive of paternalism and its wrongness with the relevant beliefs or judgements.

And so I tentatively conclude that the negative judgement about others' likely behaviour cannot explain the *pro tanto* wrongness of paternalism or indeed its distinctive nature. The initial evidentialist objection from §II is vindicated, at least *vis-à-vis* the ways of rejecting it discussed in these last two sections. The judgements about others' likely behaviour are governed by epistemic norms, and even if those are somewhat morally sensitive, still when the evidence (sufficiently) supports the judgement that someone is likely to deliberate irrationally, or to be weak-willed, or some such, that is precisely the

thing to believe. If paternalism is typically *pro tanto* wrong, its wrongness is grounded elsewhere. In the last section, I'm going to quickly suggest where. But before doing this, I want to address a worry that's already relevant here.

V

But What Do We Most Care About? There is, I think, something deeply unsatisfying about the emerging picture. The wrongness of paternalism, according to it, is about *actions*, not about *beliefs*. The relevant beliefs are either supported by the evidence or not; if not, they can be objected to on epistemic grounds; if supported, however, they cannot be objected to at all. The unsatisfying thing about this picture is that there's an important mismatch between it and some of our deepest concerns.

To see this, think again about Quong's example of your friend's request that you give him £50. Now, suppose that on the level of action, nothing paternalistic is going on at all—because of your aversion to paternalism, or for some other reason, you do give him the £50. But you're doing this in spite of your belief that he is very likely to misuse the money. And this belief of yours is, let us suppose, sufficiently supported by the evidence (perhaps according to the standards made relevant by the moral stakes involved) and indeed true. On the picture I've been arguing for, then, there is really nothing for your friend to complain about—your beliefs, governed as they are by epistemic norms, are precisely the ones called for; and your action in no way offends him, for you *do* give him the money he has asked for. And yet your friend is likely to be less than fully happy with your attitudes towards him. Perhaps, he may say, you are not exactly paternalizing him here, but your belief—that he will most likely misuse the money—is at least *patronizing*.²⁶ And this too is something he seems entitled to resent.

As is clear in this case, then, we often care deeply about what others believe about us. Indeed, we often care *more* about that than

²⁶ A fuller discussion than I can have here would include the distinction between cases in which my friend and I differ in our factual assessment of what it is that he is likely to do with the money (but agree, for each possibility, whether it amounts to misusing it), and cases in which we agree on what it is he will be doing with the money, but differ on whether or not this is a good use. (I thank Debbie Hellman for related comments.)

we do about how they behave towards us. Think again about Stroud's examples of friendship and epistemic partiality: your friend may, of course, want you not to act on a belief that he has acted shamefully. But even if no such action is relevant, it may very well be deeply important to him that you not *have* that belief. So the thought that in the relevant case there's nothing morally wrong—not even *pro tanto*—with an (evidence-supported) belief creates a mismatch between moral norms and what we care most deeply about. Similarly in the paternalism case—perhaps your friend doesn't care so much about getting the £50—not nearly as much as he cares about you not having the patronizing belief that he is likely to misuse the money. And yet if what I said in previous sections about paternalism is true, this belief of yours, if supported by the evidence, is immune to moral criticism. What is wrong with paternalism may thus be utterly divorced from what the paternalized often resent most about being paternalized.

At the end of the day—and this section—I will suggest that this is a result we're just going to have to live with. But let me first note a point that makes this result somewhat less troubling. While the concern with what others think about us is phenomenologically robust, it's not clear that it is fully coherent—and if it is, it's not clear that it's fully coherent to treat it as morally relevant or reason-giving. To see this, think again about the friend from Stroud's example—damning evidence about his shameful behaviour starts to accumulate, and it seems that he can quite sensibly care not just about what you do about it, but also about what you believe of him. But what *exactly* does he care most about here? At least in typical cases, I now want to suggest,²⁷ he doesn't merely care about your believing that he acted well (or your not believing that he acted shamefully). If you did all that simply by ignoring the evidence, he would not, I think, be entirely satisfied. What he really wants, I submit, is that you believe well of him *based on the evidence*.²⁸ We can even imagine him saying something along these lines: 'I don't want you to just take my word for it, or to ignore the evidence out of loyalty. What I really want is for you to have a hard look at the evidence, and then see that I did not act shamefully!' Again, on phenomenological grounds, this seems rather robust. But at this point it's not clear how this can be

²⁷ I don't know of anyone making quite this point in the literature. But there are related ideas in Kawall (2013, pp. 358 ff.).

²⁸ Here I include in the evidence testimonial evidence, of course.

made rational: if your friend wants you to believe well of him *based on the evidence*, if this is what he cares about, then what he seems to want is that *the evidence support* the belief that he did not act shamefully. And while it makes sense to have an attitude of this kind, this attitude is not one that is directly about you and your friendship; nor is it clear that it is a practical attitude at all. Rather, at this point it looks like a *mere wish*—the practically irrelevant, motivationally incompetent relative of a desire or a practical concern.

I don't want to overstate the relevance of this point. First, while what I just said seems to me true about central cases of caring about what others think of us, I do not, of course, claim that this is always or necessarily the case. I'm sure there are times where—perhaps knowing full well that I did behave shamefully, but desperately wanting my friend's continuing respect and love—I want his beliefs about me not to be based on the evidence. It's just that such cases seem to me atypical in an important sense: they do not represent in the best rational light the way in which we care about what others think of us. And second, I don't think there's anything necessarily irrational about mere wishes. I may wish I were the smartest person in my department, and this wish makes perfect sense, and is not, it seems to me, in itself rationally criticizable (as other wishes presumably are). It's just that wishes do not seem to have the kind of practical rational relevance that other attitudes have.

And this brings us back to the paternalism case. Perhaps what your friend cares most deeply about is not whether or not you give him the money, but whether or not you believe that if you do he will misuse it. But if the phenomenological characterization of the central case above is right, then what your friend wants is not your blind loyalty, but rather your believing *on the evidence* that he is likely to use the money well. That is, he wants the evidence to support this judgement. But now, how does this give you any relevant reason?²⁹ Either the evidence (available to you) does, or it does not, support that judgement. You can share your friend's wishes here—it will (or would) be nicer if the evidence reflected well on him. And if it doesn't, this is something to regret. But none of this seems practically relevant in the right way—none of this has anything to do, it seems,

²⁹ Getting back to a point made earlier, it may give you a reason to *look for* more exonerating evidence, and so on. But this is not a relevant point here.

with whether or not you should give him the money, or with the moral status of the paternalizing action or decision.

So the typical *pro tanto* wrongness of paternalism is not grounded in the negative belief or judgement about the paternalized's likely behaviour. And yes, this means that in many cases what makes paternalism wrong will not be the thing about it most resented, or at least regretted, by the paternalized. But the concern this resentment is tied to—the concern about what others believe of us—is in the typical case just a concern about what the evidence supports, and so more a wish than anything more practically relevant than this. So the gap between what makes paternalism wrong (when it is) and the concerns of the paternalized (in cases of interpersonal paternalism) is a gap we just have to accept.³⁰

VI

When It's Wrong to Act on Epistemically Justified Judgements: The Value of Autonomy. What grounds or explains the *pro tanto* wrongness of paternalism is, then, not something about a problematic belief that underlies paternalistic interventions, but rather something practical, something about *action*. But this does not mean that the belief (roughly, that the paternalized is unlikely to behave rationally) has no role to play here. It may even play a role in explaining the *pro tanto* wrongness of the paternalistic intervention. But the nature of this role is different from that discussed so far. The problem lies not with the belief (when it is best supported by the evidence), but rather with the action based on it. Let me explain.

At least in standard cases of paternalistic intervention, a belief that the paternalized is unlikely to act rationally will indeed be present; otherwise, if the paternalized is, say, at least as likely as the paternalizer to make the right decision (for the paternalized) and then follow through on it, why would the paternalizer think her paternalistic

³⁰ There are other examples that share this structure. Perhaps I speak to you (or in some other way act towards you) in a way that discloses absence of warmth that was once there. Perhaps my action, for whatever reason, is wrong. But perhaps what you regret most about it is that we are no longer as close as we once were. Surely, though, this can't be what makes my action wrong. And indeed, at this point, your relevant concern looks like a mere wish (though perhaps an intense one)—the wish that we hadn't drifted apart, or some such.

intervention is justified at all?³¹ So the negative judgement about the paternalized is present, and furthermore, we are assuming, that judgement is best supported by the evidence. And yet it is not the judgement that is morally (or otherwise) problematic here. What is problematic is *acting on* the judgement, or being motivated by it. The relevant judgements—that your friend will misuse the money, that we will all under-save for the far future, that I will eat too much of the candy—may all be true, and best supported by the evidence. Even if they are, though, we have moral reasons not to act on them.³² And this may seem puzzling: that your friend is likely to misuse the money seems to be normatively relevant, it seems to count against giving him the money. And if you know that he is likely to misuse the money, why not act accordingly? Why not act for the reason that he is likely to misuse the money, if need be, by refusing to give him the money?

Without pretending to offer a full account, let me make the following two points, one structural, one substantive.

First, then, regarding normative structure. The most plausible way of thinking about a situation such as this is, I think, utilizing Joseph Raz's understanding of *exclusionary reasons*.³³ Exclusionary reasons are reasons not to act on some other reasons. Exclusionary reasons *exclude* reasons, they do not *defeat*, *cancel* or *undercut* them. If, for instance, you promised to meet me for lunch, but on your way you come across someone whose company you're going to enjoy more, that she is so much better company is most definitely a reason to go with her to lunch (rather than with me), and genuinely counts in

³¹ This is simplistic in ways that need not concern us here (and hence the qualification to standard cases in the text). For instance, the paternalizer may have no beliefs at all about the paternalized's likely behaviour, and not care, either. She may just have a very high estimate of *her own* rationality, and settle for that. In such a case, her intervention will be paternalistic, and indeed typically *pro tanto* wrong in exactly the way characteristic of paternalism. Notice, by the way, that this example alone suffices to refute Quong's official characterization of paternalism, and perhaps some of the other ones I started with as well.

³² Though here the emphasis shifted from problems with the belief or judgement to problems with the action, it's not clear that what I say here is inconsistent with the accounts of paternalism I started with. Quong's relevant condition, for instance, starts thus: '[The paternalizing] A's act is motivated by a negative judgement about [the paternalized] B's ability . . .' (2011, p. 80). So it's consistent with his official account that what's wrong here is not the judgement itself, but being motivated by it. Throughout his discussion, though, Quong emphasizes the belief and its status, not being motivated by it. And he nowhere hints at what reason we may have not to act on the relevant (true, justified) belief.

³³ See Raz (1990, p. 39). I explain this mechanism in some detail in Enoch (2014, pp. 318 ff.).

favour of so doing. It's just that the promissory obligation *excludes* that reason: it provides you not just with a reason to meet me for lunch, but also with a reason not to act for the reason that you're going to enjoy lunch with her more. And this, it seems to me, is the most natural thing to say about paternalism as well. That your friend will—as you know—misuse the money gives you a reason not to give him the money (or perhaps cancels some of the reasons you otherwise would have had to give him the money); it's just that this reason is excluded—you have a reason not to act on that reason, not to give it weight in your practical deliberation, and so on. Understanding this as an exclusionary reason helps to explain why your friend is less likely to engage you in discussion on the merits of his likely use of the money, and more likely to insist that this is not even the kind of consideration you should be thinking about here. Similarly, that people will under-save for their retirement genuinely counts in favour of nudging them into giving more (it even counts in favour of taking some of their money, and saving it for them); the controversy over whether doing so would be objectionably paternalistic is best understood, I submit, as a controversy over the question of whether there is a reason for the state not to act for this reason. And so on.³⁴

But this was just structure—the substance has to be filled in. We understand why you have a reason not to act on the reason that you're going to enjoy lunch with her more, because we know what the exclusionary reason is: it's that you promised to have lunch with me. Using the device of exclusionary reasons, we understand how it's *possible* for you to have, as you know, a reason not to give your friend the money, and yet to have a moral reason not to act for that reason; but in order to see that this is actually so, what we need is an understanding of the exclusionary reason itself or the value associated with it. Similarly, the thing to do in order to promote the discussion of whether a paternalistic intervention is called for in the saving-for-retirement case is to think about the substantive values at stake, and see whether they ground a reason not to act (or perhaps not to act in certain specific ways) for the reason that is given by people's tendency to under-save.

As the geography of this paper makes clear, I won't conduct this full discussion here. But, I now want to suggest, there are not going

³⁴ The idea of an exclusionary reason is central to how Groll (2012) characterizes paternalism.

to be any huge surprises here. The relevant value, it seems to me, is the value of personal autonomy, and the questions about paternalism are best understood as questions about how best to understand that value and the reasons and obligations associated with it.³⁵ The fact that paternalism arises both in interpersonal and in political contexts is a consequence of the fact that the value of personal autonomy is relevant in both; the fact that the moral constraints regarding paternalistic interventions may be different in interpersonal and in political contexts is a consequence of the fact that personal autonomy may be significant in somewhat different ways in these contexts, and may entail different constraints.

Notice that this way of thinking about paternalism and what grounds its *pro tanto* wrongness fits very well the intuitive thought that the paternalizer engages in what is not his or her business.³⁶ When your friend claims, for instance, that whether or not he's going to misuse the money is none of your business, what he is in effect saying is that this is not a consideration you should be acting for, or even deliberating on. He is asserting his autonomy. And this shows, *pace* Quong, that the mere fact that the money is yours, not his, doesn't show that it's not his autonomy which is at stake. The value of his autonomy gives you a reason—an exclusionary reason—not to refrain from giving him the money for the reason that he is likely to misuse it. If you do refrain from giving him the money for that reason, you are in violation of his autonomy. And note that because the reason his autonomy gives you is merely the reason not to refrain from giving him the money *for the reason that he's going to misuse it*, if you end up refusing to give him the money for some other reason—such as that you want to spend it yourself—you're not in violation of his autonomy, though of course you may be in violation of some other important moral considerations.

The fact that someone is likely to act in a way that will be bad for him or her, or irrationally, or some such, does often give others relevant reasons for action. Certainly, the evidence often gives them a reason to believe that this is so. And when it does, they also *know* that they have a relevant reason for action, a reason to engage in what will

³⁵ Mostly for reasons related to his larger public-reason project, Quong (2011) resists relying on the value of autonomy here. But of course, in a general discussion of paternalism we need not be committed to Quong's larger project.

³⁶ Perhaps this is the intuition Shiffrin (2000) tries to capture with talk of having an effect on the paternalized's sphere of agency.

amount to a paternalistic intervention. But the value of the potentially paternalized's autonomy also gives a reason not to act for that first-order reason. And we can make progress on the discussion of paternalism—how best to understand it, perhaps, and much more importantly, why and when it is objectionable—by better understanding the value of autonomy and its related constraints, not by focusing attention on the beliefs about people's expected irrationality or akrasia. Such beliefs, after all, may very well be the thing to believe.³⁷

*Department of Philosophy
Hebrew University of Jerusalem
Mount Scopus Campus
Jerusalem 91905
Israel*

REFERENCES

- Bullock, Emma C. 2015: 'A Normatively Neutral Definition of Paternalism'. *Philosophical Quarterly*, 65, pp. 1–21.
- Conee, Earl, and Richard Feldman 2004: *Evidentialism: Essays in Epistemology*. Oxford: Clarendon Press.
- de Marneffe, Peter 2006: 'Avoiding Paternalism'. *Philosophy and Public Affairs*, 34, pp. 68–94.
- Driver, Julia 1999: 'Modesty and Ignorance'. *Ethics*, 109, pp. 827–34.
- Dworkin, Gerald 2013: 'Defining Paternalism'. In Christian Coons and Michael Weber (eds.), *Paternalism: Theory and Practice*, pp. 25–38. Cambridge: Cambridge University Press.
- Enoch, David 2006: 'Taking Disagreement Seriously: On Jeremy Waldron's *Law and Disagreement*'. *Israel Law Review*, 39 (3), pp. 22–35.
- 2014: 'Authority and Reason-Giving'. *Philosophy and Phenomenological Research*, 89, pp. 296–332.
- forthcoming: 'Political Philosophy and Epistemology: The Case of Public Reason'. To appear in *Oxford Studies in Political Philosophy*.
- and Talia Fisher 2015: 'Sense and Sensitivity'. *Stanford Law Review*, 67, pp. 557–611.

³⁷ I presented earlier versions of this paper at the University of Virginia School of Law, Columbia Law School, Brooklyn Law School, and at the meeting of the Aristotelian Society. I thank the audiences on those occasions for their helpful discussions and comments. I also thank Anne Coughlin, Adam Colbert, Hanoch Dagan, Amanda Greene, David Heyd, Susan James, Lucy O'Brien, David Owens, Alex Raskolnikov and Matthew Soteriou. Special thanks to Joey van Weelden for early help with references.

- Levi Spectre, and Talia Fisher 2012: 'Statistical Evidence, Sensitivity, and the Legal Value of Knowledge'. *Philosophy and Public Affairs*, 40, pp. 197–224.
- Fricker, Miranda 2007: *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.
- Groll, Daniel 2012: 'Paternalism, Respect, and the Will'. *Ethics*, 122, pp. 692–720.
- Hawthorne, John 2004: *Knowledge and Lotteries*. Oxford: Clarendon Press.
- Ichikawa, Jonathan Jenkins, and Matthias Steup 2012: 'The Analysis of Knowledge'. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/knowledge-analysis/>.
- Kawall, Jason 2013: 'Friendship and Epistemic Norms'. *Philosophical Studies*, 165 (2), pp. 349–70.
- Keller, Simon 2004: 'Friendship and Belief'. *Philosophical Papers*, 33 (3), pp. 329–51.
- 2007: *The Limits of Loyalty*. Cambridge: Cambridge University Press.
- McGeer, Victoria 2008: 'Trust, Hope, and Empowerment'. *Australasian Journal of Philosophy*, 86, pp. 237–54.
- Mill, John Stuart 1859: *On Liberty*. Available online at *Project Gutenberg*. <https://www.gutenberg.org/files/34901/34901-h/34901-h.htm>.
- Mittag, Daniel M. 2004: 'Evidentialism'. In *The Internet Encyclopedia of Philosophy*. <http://www.iep.utm.edu/evidenti/>.
- Nagel, Thomas Nagel 1987: 'Moral Conflict and Political Legitimacy'. *Philosophy and Public Affairs*, 26, pp. 215–40.
- Pace, Michael 2011: 'The Epistemic Value of Moral Considerations: Justification, Moral Encroachment, and James's "Will to Believe"'. *Noûs*, 45, pp. 239–68.
- Quong, Jonathan 2011: *Liberalism Without Perfectionism*. Oxford: Oxford University Press.
- Raz, Joseph 1990: *Practical Reason and Norms*. Princeton, NJ: Princeton University Press.
- Shiffrin, Seana 2000: 'Paternalism, Unconscionability Doctrine, and Accommodation'. *Philosophy and Public Affairs*, 29, pp. 205–50.
- Stanley, Jason 2005: *Knowledge and Practical Interests*. New York: Oxford University Press.
- Stroud, Sarah 2006: 'Epistemic Partiality in Friendship'. *Ethics*, 116, pp. 498–524.
- Thaler, Richard H., and Cass R. Sunstein 2008: *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
- Tsai, George 2014: 'Rational Persuasion as Paternalism'. *Philosophy and Public Affairs*, 42, pp. 78–112.
- Waldron, Jeremy 1999: *Law and Disagreement*. Oxford: Oxford University Press.