

Feeling Pain for the Very First Time: The Normative Knowledge Argument

Guy Kahane, University of Oxford

Abstract. In this paper I present a new argument against internalist theories of practical reason. My argument is inspired by Frank Jackson's celebrated Knowledge Argument. I ask what will happen when an agent experiences pain for the first time. Such an agent, I argue, will gain new normative knowledge that internalism cannot explain. This argument presents a similar difficulty for other subjectivist and constructivist theories of practical reason and value. I end by suggesting that some debates in meta-ethics and in the philosophy of mind might be more closely intertwined than philosophers in either area would like to believe.

According to internalist theories of practical reason, what we have reason to do is whatever would fulfil either our actual present desires or the desires we would have had we gone through some procedure of informed deliberation whose starting point is our 'motivational set'.¹ On opposing externalist theories, what we have reason to do is not determined by our present beliefs and desires and what can be reached from them by some deliberative route. Many philosophers believe that externalist theories commit us to metaphysically queer normative facts and epistemic powers that don't fit a naturalistic view of the world. This is one reason why externalism is often rejected.

In this paper, I present a new argument against internalist theories. I use 'internalism' in the broadest sense, to refer not only to Bernard Williams's view, but to a wide range of desire-based or subjectivist views, ranging from the simplest Humean instrumentalism all the way to elaborate forms of Kantian constructivism.² And although the target of my argument are internalist views of

¹ Bernard Williams first introduced internalism in Williams 1981. Williams takes our 'subjective motivational set' to include not only desires but also dispositions of evaluation, patterns of emotional reaction, and personal and projects (1981, 105). But in what follows I shall use 'desire' in the broad sense that includes these and other conative attitudes.

² Internalism in the sense I shall be discussing shouldn't be confused with what is sometimes called 'judgement internalism' – a view about the relation between moral or evaluative *beliefs* and motivation. The view I am discussing is a view about the relation between *practical reasons* and motivation. I will assume that internalism in this sense states both necessary and sufficient conditions for the existence of a reason. Williams sometimes presents his own brand of internalism as stating only a necessary condition, although he admits that he also takes it to be sufficient (see Williams 1989, 39, 35). Such a weaker view wouldn't relieve us from the threat of queer normative facts: it is compatible with thinking that such queer properties give us reasons only on the further condition that they motivated us. My target here is thus the stronger view. But some of my arguments also have force against the weaker view. I'll highlight this when

practical reason, it can be extended to target parallel subjectivist or dispositional views of value.

I call this argument the Normative Knowledge Argument. As will soon be apparent, it is directly inspired by Frank Jackson's celebrated Knowledge Argument.³ My argument, however, is meant to be a contribution to debates about practical reason and value, not to debates about qualia and physicalism. As will emerge, however, these two debates may not be entirely separable. I shall only have space to identify some broad connections.

I shall proceed as follows. I shall start by examining a challenge that Williams has set to externalist views. This challenge, when properly interpreted, has not yet been adequately answered. I shall then develop a thought experiment that is meant to answer this challenge. The challenge will not only be answered. It will be reversed.

WILLIAMS'S CHALLENGE

I.

In a crucial passage, Williams sets the debate between internalists and externalists as follows. He notes that whenever an agent acts for a reason, this reason could always be truly described as an internal reason. It does not follow, of course, that there are no external reasons. What does follow, Williams suggests, is that the content of external reason statements

is not going to be revealed by considering merely the state of someone who believes such a statement, for that state is merely the state with regard to which an internal reason statement could be truly made. Rather, the content of the external type of statement will have to be revealed by considering what it is to *come to believe* such a statement.⁴

In a later paper, Williams challenges externalists to explain what is it, on their view, that

the agent comes to believe when he comes to believe he has a reason to Φ ? If he becomes persuaded of this supposedly external truth, so that the reason does then enter his motivational set, what is it that he has come to believe? The question presents a challenge to the externalist theorist.⁵

These passages have sometimes been read as asking for a *reductive analysis*

appropriate.

³ The knowledge argument was introduced in Jackson 1982. The subsequent literature is voluminous. For a survey of the current state of the debate, see Ludlow, Nagasawa and Stoljar 2004.

⁴ Williams 1982, 107-8.

⁵ Williams 1989, 39.

of the notion of a reason for action in terms of other, presumably non-normative notions. If this was all there is to the challenge, then externalists can legitimately decline to answer it. They can deny that the concept of a normative reason can be explained in terms of other, more basic concepts. They can claim that normative concepts are irreducibly basic.⁶ Moreover, unless internalism itself takes such a reductive form, then internal reason statements themselves make use of the very same normative notions. So if the challenge is to be understood in this way, it is not clear why it is addressed specifically to the externalist. On another reading, what Williams is asking is an explanation of how a recalcitrant agent could *rationally arrive* at a new normative belief if not by some deliberative route from his existing motivational set. This again seems to beg the question. For why can't agents arrive at a new normative belief by conversion, or by rational insight?⁷

II.

In one of Williams's examples, we are presented with a husband who badly mistreats his wife. This husband, we are told, is simply incapable of being motivated to behave more kindly, and this means that, whatever else we may want to say about him, we cannot say that he has an internal reason to behave better.⁸ Suppose, though, that at some later point this husband changes his ways. It is easy to see what internalists can say about what happened: there was some non-rational change in the husband's motivational set, and he consequently acquired an internal reason to behave kindly, a reason that wasn't there beforehand. Williams challenges externalists to explain what else might have happened here. And we saw what externalists reply: that what the abusive husband came to believe is simply that he has – and always had – a reason to treat his wife more kindly, and he came to believe that by *seeing* that this is what he has reason to do. Nothing more needs to be said.

This reply dismisses Williams's challenge rather briskly. All that externalists have shown, however, is that they can give an alternative explanation of what has happened. But while sometimes Williams seems to doubt that externalist reason claims even have a clear *sense*, at other times what he doubts is rather that there are any grounds for thinking that any such claim is *true*. This latter thought allows us to restate Williams's challenge to externalism. It is not that externalists cannot explain changes in normative belief. It is rather that they have not given us any reason to prefer their explanation to the internalist one.

Williams starts with the point that whenever an agent does acknowledge and act for a reason, we can always describe this agent as acknowledging and

⁶ For this reply, see Parfit 1997, 121, and Gibbard 2003.

⁷ For appeal to conversion, see McDowell, 1995. For appeal to rational insight, see Parfit, 1997, 118.

⁸ Williams 1989, 39-40.

acting on an *internal* reason. What work, then, is there for external reasons to do? Do we really need to postulate their existence to best explain normative thought and discourse? There are a number of ways in which the internalist explanation of this discourse has the explanatory advantage. Most familiarly, the internalist conception can easily explain the conceptual tie between normative belief and motivation and action.⁹ But internalists can also explain how people change their normative views simply by pointing out changes in their motivational set, and the causes for these changes. They can explain the epistemology of normative belief using familiar materials from the philosophy of mind: our first-person authority over our intentional states, including our desires, and our capacity for logical and causal reasoning. They can also easily explain an agent's *failure* to know what he has reason to do, by citing the various ways in which people sometimes fail to know what they really want, or the ways in which their reasoning may slip. If there is disagreement over the existence of an *internal* reason, there will always be an explanation of why one of the parties got things wrong. To put things politely, externalists have rather less to say about all of these.

On this reading, Williams isn't claiming that external reason statements are incoherent or that internalism is the true account of the meaning of everyday statements about reasons.¹⁰ Williams is best understood to be raising an *explanatory* challenge. What he is claiming is that internalism gives a more plausible explanation of the phenomena. The challenge to externalists is to explain why the internalist explanation is not good enough.

The externalist claims that when the abusive husband mends his ways, he does so because he has come to believe that he has a reason to behave better, and has come to realise that through conversion or insight. But Williams could still press the challenge. What Williams needs to say is simply that the externalist explanation of the change is *redundant*. It is redundant because we can already give a complete explanation of what happened in internalist terms, as involving a change in the husband's motivational set. Williams can next claim that his explanation does not merely win on points, by standard criteria of good explanation, but *decisively*, because the externalist explanation requires appeal to *queer normative facts and cognitive powers*, whereas the internalist explanation needs nothing more than uncontroversial materials drawn from familiar philosophy of mind.¹¹ So why think that anything else is at work here?

⁹ Some of the things Williams says support this explanatory reading. For example, in Williams 1989, he raises his challenge to externalists after noting that the internalist conception is at an advantage because it's able to explain how an agent's belief that he has a certain reason for action could lead him to act appropriately. "It is obvious on the internalist view," he writes, "how this works."

¹⁰ Williams sometimes writes as if internalism gives the sense of statements about reasons (see e.g. Williams 1989, 40), but he has elsewhere explicitly denied that this is how internalism is best understood (cf. Williams 1995, 188).

¹¹ As I noted above, a non-reductive form of internalism will also be open to a charge of

In this paper I aim to meet Williams's challenge. I have presented it as having two parts: the externalist explanation of apparent change in normative view is claimed to be at once redundant and queer. These are separate claims, since an explanation can be implausible or redundant without being queer. But presumably if we can give a satisfying explanation of something using familiar materials while the alternative explanation appeals to something queer, this would further support the a charge of redundancy. In what follows I shall address only the first part of the challenge. I shall show that a certain example of change in normative belief can *only* be explained in externalist terms. The internalist alternative is simply inadequate. This does not mean that what happens when an agent changes his normative views may not be philosophically queer. But not everything queer is redundant.

My argument for this conclusion will make use of an example that is rather fantastic. What I shall ask is what happens when an agent experiences pain for the first time. I'll argue that such an agent will gain new normative knowledge that internalism cannot explain.

BEFORE PAIN

I.

I shall assume two unremarkable claims about pain. One is an evaluative claim. I take it that

(A) Pain is intrinsically bad for the person experiencing it

The other is a normative claim, the claim that

(B) Pain gives the sufferer *pro tanto* agent-relative reasons to prevent, end or minimise it

There are different ways of understanding the relation between (A) and (B). On a buck-passing view, (A) is to be understood in terms of (B): for pain to be intrinsically bad simply is for it to be such that it gives reasons to want to avoid it.¹² On the opposing view, (A) states a prior evaluative fact that *explains* why (B) is true. I need not adjudicate between these two views about the relation between value and normative reasons. But given the close connection between

queerness, even if a somewhat weaker one.

¹² See Scanlon 1998, 95-100. If the buck-passing account of value is correct, then my argument will immediately also apply to subjectivist and dispositional theories of value. I believe it can be fairly easily adjusted to target such theories of value even if the buck-passing account is false, but I do not have space here to directly discuss such theories.

the two on either view, I will allow myself to speak interchangeably of pain's badness, and of the reasons provided by pain.

II.

Premise (1) of the argument is a modal claim. What I ask you to conceive is someone – call him Zeno – who has never felt pain.

There is no difficulty in Zeno's never having felt *physical* pain. There *are* people who are incapable of feeling physical pain.¹³ These actual cases, however, might be too weak. Even someone who has never felt physical pain may still be familiar with other forms of unpleasantness, or with physical pleasure. And, for reasons I shall explain below, we may need our imaginary character to be utterly innocent when it comes to hedonic experience. In order not to beg any questions, I shall adopt here the widest understanding of the hedonic, taking it to include not only all unpleasant physical experiences, but also so-called mental pains, and all the corresponding positive or pleasant experiences.

To conceive of someone who meets *this* condition is certainly harder.¹⁴ I shall not say much at this point to establish that such hedonic innocence is possible. I know of no good arguments *against* its possibility. Furthermore, even if it may be hard to conceive of such a state, this difficulty has nothing to do with the dispute between internalists and externalists. There is nothing in internalism that rules out the possibility of my imaginary scenario.

One barrier to admitting the possibility of hedonic innocence, I think, has to do not with conceiving a *state* of complete hedonic innocence, but of conceiving a *life* from which the hedonic is entirely absent. If this is a genuine difficulty, then it can be easily addressed by weakening the scenario. We can suppose that although Zeno's past was as riddled with hedonic incident as ours, all trace of that has been washed away from his memory. He cannot remember ever having felt pain or pleasure. He cannot remember what that would feel like. Even if we cannot imagine a life without hedonic incident, we can certainly imagine hours or days like that. And this is enough for us to pose our question. What matters is that Zeno has no *access* to how the hedonic feels like – no access through experience, memory or imagination.

III.

My first premise was that it is possible for someone not to have access to what

¹³ See Nagasako, Oaklander, & Dworkin 2003.

¹⁴ How could Zeno be entirely immune to pain and other hedonic states? The details don't matter much. We can imagine Zeno to be congenitally insensitive to all kinds of positive and aversive bodily sensations. Or we can imagine him to have anaesthesia continuously running through his blood.

the hedonic feels like. Since the argument is going to revolve around claims about knowledge, let me make clear what Zeno *does* know in his innocent state. For simplicity's sake I shall focus on pain but what I say should be taken to refer to everything hedonic.

Assume that Zeno is a rational agent who possesses the concept of an experiential state and basic evaluative and normative concepts such as those of badness and of a normative reason. What Zeno knows is that other people experience something they call pain. Suppose also that after much study, Zeno has become an expert on pain. He knows

Pain's functional properties. Zeno knows pain's functional role – its normal causal antecedents and consequents. He knows, for example, that contact with fire normally causes pain, and that to feel pain makes you withdraw your arm and cry out. Most importantly, he knows its *motivational* properties: he knows that when people experience pain, they dislike it and want it to end.

Its role in our social practice. Zeno can also know the role that the state of feeling pain plays in our practice. He knows, for example, the attitudes people normally have towards pain: that they fear it and are much relieved when it's over, and that people resent and treat as morally wrong to gratuitously inflict pain on others.

Zeno knows exactly *how* he will react when he first experiences pain. What he doesn't seem to know is *why*. And although he is familiar with the complex social practice that surrounds pain, it seems to me that Zeno still does not know *why* our practice is shaped in this way: what *justifies* and *makes sense* this complex pattern of attitude and belief.

What Zeno knows are some of pain's non-normative properties. But it is at least plausible that by knowing these he will also acquire some normative information about it:

Indirect normative information. Zeno would presumably know that we *believe* pain to be bad.

In fact Zeno could even take it, on testimony, that pain *is* indeed bad. That is, he could accept this proposition as true. But it still seems that Zeno is lacking something vital: he still doesn't know *why* pain is bad. To *accept* that the propositions 'pain is bad' and 'pain gives reasons to try to prevent it' are true is not the same as to know what *makes* them true.¹⁵

¹⁵ I do not need to take a stand here on whether Zeno could be said to *know* that pain is bad on the basis of testimony without, however, really understanding *why* it's bad. In this loose sense, a congenitally blind person could also be said to know that Velasquez's *Las Meninas* is beautiful.

IV.

A further premise of the argument is nearly explicit in the above remarks: that there is something normative that Zeno wouldn't know about pain which he would know only when he first experiences it. This claim is based on intuition: the intuition that Zeno is lacking a crucial piece of *normative knowledge*. Call this the *normative knowledge intuition*. And it seems that Zeno is lacking this knowledge *because* he does not know what it's like to suffer – because he lacks a crucial bit of *phenomenal knowledge*. Someone who doesn't know what it's like to suffer couldn't know why suffering is bad, why we are justified in fearing it or in trying to prevent it.

I can now explain why I've needed Zeno to be in a state of *complete* hedonic innocence: it's hard to rule out the possibility that he could come to know that pain is bad on the basis of his knowledge that, say, pleasure is good. After all, even if I have never felt seasickness and thus don't know what it's like to be seasick, I can still understand what it means to say that seasickness feels bad – feels bad in the very same way that headaches and burns feel bad. But although it is hard to rule out this possibility, it is far from obvious that there actually is such an intimate tie between the epistemology of pain's badness and that of pleasure and other hedonic states. I am thus making a very significant concession to the internalist in requiring Zeno to be hedonically innocent.

The normative knowledge intuition thus involves a claim about missing phenomenal knowledge:

(2) Zeno wouldn't *know* what it's like to feel pain

I take it to be uncontroversial that if Zeno has never experienced pain and has no access to this experience through memory or imagination, then he doesn't know what it's like. What it is exactly that he doesn't know depends on our understanding of pain's badness. On a natural view, *the Sensation Theory*, the sensation of pain is intrinsically bad – its badness has nothing to do with whether or we dislike it or not. On this view

(2a) Zeno wouldn't know what it's like to feel the sensation of pain

This view is not compatible with internalism. If it were true that the sensation of pain is intrinsically bad, and if, as we have been assuming, pain's badness implied the existence of reasons, then pain would give us reasons to avoid it quite independently of whether or not we dislike it. These would have to be external reasons.

Many philosophers, however, reject the Sensation Theory. They reject it because there seem to be cases, such as when patients undergo frontal lobotomy,

where the sensation of pain seems to lose its hurtfulness. When lobotomy patients can still feel the sensation of pain but they do not dislike or mind it, and there seems to be no reason to suppose that what they experience is nevertheless bad. Most philosophers consequently hold a version of what I shall call *the Dislike Theory*. On this view, whenever pain is intrinsically bad, it is also necessarily disliked – necessarily the object of a negative attitude. In this paper, I shall assume the truth of the Dislike Theory. I assume it both because it is independently plausible, and because if the Sensation Theory were true then we could reject internalism without further argument.¹⁶

As the Dislike Theory is usually understood, it is the claim that to suffer is simply to be in a state of wanting what is in itself a neutral sensation to stop. Pain is bad only because we want it to stop, and its phenomenal character plays no role in making it bad¹⁷ – a view of pain that certainly looks congenial to internalism and that has been understandably taken to support subjectivist accounts of value. It is sometimes assumed that if the Dislike Theory is correct, then an objectivist account of pain’s badness has to be false. As we’ll soon see, this is mistaken. What is correct, however, is that internalism and other subjectivist views are committed to *this* understanding of the Dislike Theory. I’ll call it the *Pure Dislike Theory*.¹⁸

According to all versions of the Dislike Theory, the sensation of pain is a neutral sensation that, as a matter of contingent fact, we all happen to dislike. But if the sensation of pain is a separable neutral state, then let us also assume that among the things that Zeno knows is what it’s like to feel *this* sensation. The experience he never had is rather that of *disliking* a sensation.¹⁹

We therefore need to be careful to distinguish claims about the neutral sensation of pain from claims about the state of disliking this sensation. The ordinary word ‘pain’ can be used to refer to both states.²⁰ This can be a source of confusion. To avoid such confusion, I shall use ‘pain’, ‘painful’ and ‘suffering’ to refer only to the composite mental state. My earlier description of Zeno and what he knows is to be understood in this way. When I want to refer to the neutral sensation, I shall speak of the *sensation of pain*. The sensation of pain is the

¹⁶ I discuss these theories of pain’s badness in greater detail in Kahane (forthcoming).

¹⁷ This is pretty much the philosophical consensus about pain’s badness. Let me pick just two examples from recent ethics and philosophy of mind. David Brink writes that “[P]ain is a mental state or sensation such that the person having it wants it to cease and will, *ceteris paribus*, take action to make it stop.” (Brink 1997, 112) And Austen Clark claims that “...there is no phenomenological character specific to painfulness. At best there is a phenomenological character of sensations that are painful.” (Clark 2005)

¹⁸ I’ll later discuss one unattractive way in which internalism might be compatible, in a sense, with another version of the Dislike theory.

¹⁹ Note that I am using ‘dislike’ as a technical term to refer to that distinctive affective or conative that, according to the Dislike Theory, is the source of pain’s badness. I doubt that everyday talk about the things we dislike implies the existence of such a mental state.

²⁰ This ambiguity is noted in Hare 1964.

sensation that we dislike, and that lobotomy patients don't.

With this clarification in place, we can state what Zeno wouldn't know in terms of the Dislike Theory:

(2b) Zeno wouldn't know what it's like to dislike a present sensation of pain

Think of Zeno's situation as roughly like that of the lobotomy patients, only in reverse.²¹ Just as there is a point in which a lobotomy patient fresh out of the operating room first experiences the sensation of pain unaccompanied by dislike, so Zeno, who has so far felt this sensation *only* as a neutral sensation, could at some point have his first taste of a painful experience.

THE FIRST PAIN

One day, the doctors remove the physical condition that prevents Zeno from feeling pain. What would happen when Zeno is first pricked with a sharp needle?²² For the very first time in his life, Zeno feels a sharp painful sensation. What can internalism tell us about what happens? I claim the following:

Internalism cannot explain the intuition that when Zeno first feels pain, he acquires new normative knowledge.

Internalists have to admit that there is no piece of normatively relevant information that Zeno is missing before he first feels pain. Zeno knows what the sensation of pain feels like, and he knows what it is to be in a state of wanting something to end. It's just that he happens not to have been in a state that combines the two.

It is true that in his state of hedonic innocence, Zeno cannot know that he *has* (internal) reason to avoid pain, precisely because he doesn't now want any sensation to stop, and by assumption doesn't mind it occurring in the future. So

²¹ Zeno's situation is only roughly like theirs, since lobotomy patients, possibly to a somewhat diminished degree, can still feel pleasure and suffer mentally. And it is not clear whether, even long after their operation, lobotomy patients no longer have knowledge of the source of the badness of pain.

²² Danziger and Willer recently described a fascinating case of a woman congenitally insensitive to pain who seems to have felt physical pain only once in her life. This patient reported that "the only pain she had ever felt consisted of an episode of intense headache, which had taken place 2 years before the interview but, fortunately, had been well documented by her general practitioner. This inaugural headache occurred in a context of strong emotional overload and anxiety... She had the sensation of carrying an enormous weight bilaterally on top of her head, which hurt so much that she was no more able to concentrate on her work." (Danziger & Willer 2005)

of course there isn't yet any such internal reason to know. But Zeno *does* know that when he first feels pain, he will desire a sensation to stop and will thus have an internal reason he doesn't have now. On versions of internalism on which our reasons are determined not by present motivation but by whatever motivation we have after some procedure of deliberation, Zeno would need to know that he would be motivated to end pain after the relevant deliberative procedure. But there is no difficulty in further assuming that Zeno would know that in his state of hedonic innocence.

It nevertheless seems clear that there is something of normative significance that Zeno would *discover* when he first feels pain. He would finally understand *why* it is that he has a reason to avoid pain. Internalism cannot explain this intuition.

These claims form the core of the Normative Knowledge Argument. We can state this part of the argument as follows. At t_1 , in a state of hedonic innocence,

- (a) Zeno doesn't know what it's like to dislike the sensation of pain
- (b) Zeno knows that at t_2 he will dislike the sensation of pain
- (c) Zeno is in a position to know that he will have an internal reason at t_2
- (d) According to internalism, Zeno knows all the intrinsic normative facts about his state at t_2

After pain, at t_2 ,

- (e) Zeno comes to know what it's like to dislike the sensation of pain
- (f) Zeno comes to know a new intrinsic normative fact

Therefore

- (g) Internalism is false

Both externalists and internalists can agree that when someone is in pain, he has reason to end the pain. They disagree about what explains this reason. Externalists are best understood as claiming that although Zeno can know in advance the truth of a certain claim about internal reasons, this claim doesn't refer to genuine reasons. They claim that simply by considering the functional and motivational facts, Zeno is not really in a position to know that he will have any reason to avoid pain when it first comes. I believe that our intuition about what Zeno learns from his first pain supports this claim. It is only when he first feels pain that Zeno really knows he has a reason to end pain.²³

²³ Notice that the truth of externalism is compatible with the existence of internal reasons. Externalism only requires that there exist at least *some* reasons that are not grounded in an agent's

This claim, however, might seem to beg the question. So let us suppose that Zeno is himself an internalist. He is not just *in a position* to know that he will have an internal reason, but actually knows this, on the basis of his knowledge of the motivational facts. I claim that our intuition about his first pain is still that Zeno will learn a further normative fact. It is only when Zeno first experiences pain that he understands *why* he has a reason to avoid pain. The motivational facts on their own don't provide such an explanation. And since internalism doesn't leave room for such a further explanation, what is being explained couldn't be the existence of an internal reason.²⁴

Internalism cannot explain the intuition that Zeno, before the first pain, is missing normative knowledge because he had not yet felt pain.

This is, perhaps, just another way of making the previous point. Consider Zeno before he first feels pain. Internalists are committed to holding that Zeno already knows *all* the normative facts there is to know about that future state – for what is it, on their view, that he wouldn't know? It seems, however, that if someone knew what pain is, he would *realise* he has a reason to avoid it – that he now has such a reason, even if the pain is in the future. But this is not so on the internalist conception. Although internalists must hold that Zeno already knows *all* there is to know about pain, he needn't have any (internal) reason to care about it if he doesn't happen to care about it now.²⁵ On the other hand, it seems that given what Zeno *does* know, he needn't accept that he has any reason to prevent pain. For him it should be an entirely open question whether he should care about pain, a question that will continue to remain open *until* he finally experiences pain. So it seems that there's something crucial about pain the Zeno doesn't

motivational states. It would therefore be sufficient for my purposes if what Zeno came to know was the existence of an external reason to avoid pain *in addition* to the internal one he was already aware of. Although my argument could be stated in this way, this weaker claim seems to me false to the intuitions elicited by my thought experiment. I therefore present the argument in this stronger form that denies that there is an internal reason to avoid pain, although this claim doesn't commit me to a general denial of the existence of internal reasons. I am grateful to an anonymous referee for pressing this point.

²⁴ Even externalists can allow that even in his state of hedonic innocence Zeno could, in one sense, know through testimony that he has reasons to avoid pain. If Zeno knows this then even on the externalist view what he would learn when he first feels pain is not the existence of these reasons but *why* pain is a source of such reasons. Notice however that on the internalist view Zeno has no need to rely on testimony. Since he knows the complete motivational facts, there is no further information about pain's intrinsic normative significance that testimony could reveal to him.

²⁵ Although the conjunction of internalism and the Dislike Theory of pain's badness may imply that whenever someone is in pain, he has a pro tanto reason to try to end it, the Dislike Theory doesn't not imply and cannot guarantee the existence of any internal reasons to prevent one's future pain. The Dislike Theory is best understood as a claim only about one's attitude to present sensations of pain. It is entirely compatible with complete indifference to future pain.

know when he's hedonically innocent.

Furthermore, in principle internalism allows that even before he has ever experienced pain, Zeno *could* happen to strongly desire not to be in that state. This would both give him an internal reason to prevent his future pain and make it *irrational* for him not to respond to this reason unless it is outweighed by stronger opposing reasons.²⁶ The problem is that it *doesn't* seem that it's irrational for Zeno *not* to try to prevent this future pain, given what he does know, and what he doesn't yet know about pain. That is, Zeno would *not* be open to rational criticism for failing to try to prevent the pain.²⁷ Worse, the internalist must admit that Zeno, if he does desire to avoid future pain, is in fact in the *same state* we are in when *we* care about our future pain. This again seems wrong. If Zeno had such motivation and tried to prevent his future pain, he would be just following blind motivation. We know *why* we should avoid pain.²⁸

The problem is due to the fact that from Zeno's hedonically innocent point of view, there is no difference between the state he will be in when he first suffers and a state of arbitrary compulsion. But most of us would agree that these are states with a very different normative status. I don't discover any new *normative* information when I first feel a compulsive urge. Writing of pleasure rather than pain, Quinn puts this point nicely:

Suppose I tell you that if you start scratching your ear the experience will strongly dispose you to keep on scratching. Does this by itself give you reason to want to scratch? Conceived as a kind of psychological inertial force, pleasure takes on a somewhat sinister aspect. This is because the account leaves out the salient thing: that an agent wants to prolong a pleasant experience precisely because it is pleasant – because it feels good.²⁹

²⁶ Setting aside, of course, reasons Zeno may accept on the basis of testimony. But recall that what Zeno would be accepting through testimony couldn't be any kind of further information about what *internal* reasons he has.

²⁷ Two caveats. First, this claim is compatible with the externalist claim that pain is a source of reasons to avoid it that are motivation-independent. Although there are, in a sense, reasons for Zeno to aim to avoid future pain, he is nevertheless not rationally culpable for failing to do so because he is not in a position to *know* that he has these reasons. Contrast Zeno with the abusive husband, who *should have* known better. Second, it might be objected that Zeno could still be charged with being *instrumentally* irrational if he aimed to avoid this future state yet failed to take the means to this end. Even if one granted this point, I believe that the intuition persists that Zeno wouldn't be open to the kind of rational criticism merited by someone who knew what it's like to feel pain yet failed to take appropriate means to prevent it.

²⁸ I do not see how it can help the internalist to point out that such a 'blind' desire to avoid that future state would not be *caused* by past states of such dislike, whereas our own corresponding motivation is causally dependent in this way. I fail to see why, by internalist lights, this should make any difference, let alone explain our intuitions about Zeno's epistemic state.

²⁹ Quinn 1995, 243-244.

To be sure, internalists could dismiss Quinn's question. They cannot be expected to tell us why we have reason to want to be in such a mental state, since they reject the very idea of reasons for desiring. For them, the question can only be whether we have reason to bring that future state about, and that must ultimately depend on what we happen to want now. Even if we direct our attention to the character of that future mental state, we only do so as a way of finding out what we want now for later. But there is a way of understanding Quinn's remark that doesn't beg the question against internalism. The passage could be read as claiming that, given that we *are* motivated by the prospect of future pleasure, and given that we remain cold to the prospect of the kind of state that pleasure would have to be if internalism is true, then the internalist explanation of why we actually are motivated by future pleasure couldn't be right.

Internalism cannot account for the epistemic modality of the normative knowledge that is acquired when one first feels pain.

Turn now to a different scenario on which Zeno, in his state of hedonic innocence, doesn't even know that a state such as pain even exists. A headache catches him entirely by surprise. What does Zeno come to know when he first feels hurtful pain? It seems that from the first moment, Zeno would know that he is experiencing something intrinsically bad. He would come to know with certainty that he has reason to get rid of this state. And he would come to know that this is a state he has reason to prevent and get rid of in any future instance.

Consider how internalists could try to account for this knowledge. Zeno is now for the first time in a state of disliking the sensation of pain. So what he comes to know is a psychological fact: that he is motivated to get rid of this sensation. On the basis of this psychological fact Zeno also comes to know a normative fact: that he now has an internal reason to take the means to get rid of this sensation. But there is immediately a problem about the degree of *certainty* with which Zeno could know this normative fact. This is because, for all Zeno knows, it is possible that he may be mistaken in thinking he really has a negative attitude to the sensation of pain. We have first-person authority with respect to our desires and other intentional states, but we can certainly be mistaken about what we really want. And if he has internal reason to get rid of it only if this desire would survive a certain *deliberative procedure*, then he is even more liable to be mistaken in thinking he has such a reason. So even if exposure to pain does provide Zeno with knowledge of the (subjectivist) badness of pain and of the (internal) reasons it gives him, this knowledge doesn't seem to have the right kind of epistemic modality. For, to repeat, it seems that when Zeno first feels the pain he immediately knows with certainty that this strange new experience is bad and gives reasons to get rid of it. The epistemology of the badness of suffering is like the epistemology of a self-evident truth or of an immediate

experience,³⁰ not of a propositional attitude such as desire.³¹

The internalist faces further problems with respect to our knowledge of the *degree* of suffering. If the badness of pain is to be explained by the presence of a negative conative attitude of a certain strength, it is not clear how we could ever directly know, just by experiencing an episode of pain, just how bad it is. After all the strength of a desire seems to be a *relational* property, roughly, one derived from its tendency to issue in appropriate behaviour when combined with the agent's beliefs and his other desires. But if so, it would be quite hard to explain how we could ever directly know how bad our pain is, as we evidently can.

These objections to internalism targeted Zeno's knowledge that he *now* has reason to get rid of *this* sensation. But for internalists this reason is in principle separable from any reason to prevent or get rid of pain over time. The motivation that gives this further, ongoing reason is contingent and distinct from the dislike that is, according to the Dislike Theory, a necessary component of suffering. But how, just by experiencing a sharp headache, could I know that I have such a stable and lasting motivational state – that my desire to avoid pain would extend into the far future? I could only come to know *that* with any degree of confidence on the basis of inductive evidence gathered over time. Yet it seems that the existence of such a general reason to avoid pain, and thus of whatever may underlie it, is something Zeno would know right away.³²

ANSWERING WILLIAMS'S CHALLENGE

I.

Return now to Williams's challenge. Of supposed examples of change in

³⁰ In fact I believe that the epistemology of suffering combines both: someone confronted with the immediate experience of pain is in a position to know with certainty the self-evidence of (A) and (B).

³¹ It might be objected that affective or conative attitudes can also have an experiential dimension. I can feel fear or have a felt desire to do something. This is certainly true. The problem is that having such experiences is not a sufficient condition for having the attitude they normally intimate. I can feel the pull of a desire to ϕ without really wanting to ϕ . The two kinds of states can sometimes come apart, and it is my overall pattern of behavioural dispositions which determines whether I can really be said to want to ϕ or not, not the momentary occurrent feeling. So knowledge of such feelings isn't sufficient for knowledge of the dispositional state of which they may be manifestations. But the situation is exactly the reverse when it comes to suffering. The feeling of suffering is not an occurrent manifestation, in some cases illusory, of some dispositional state. It is all there is to suffering. For discussion of this basic distinction between the epistemology of states of consciousness and that of propositional attitudes, see the introduction to Macdonald, Smith, and Wright, 1998.

³² This objection to internalism has force even against the weaker version that sees the tie between reasons and motivation as only a necessary condition (see fn. 2 above).

normative views – of, say, an abusive husband coming to think that he should stop mistreating his wife – internalists claim the following:

Before the change, the agent had no internal reason to ϕ , and ascription to him of external reasons that in no way connect with his point of view and motivation is empty and redundant.

After the change, the agent has an internal reason to ϕ . So the claim that he is responding to an external reason is again redundant.

As for *the point of change*, the externalist leaves it completely mysterious both why the agent failed to see the reason earlier, and what changed in him that allows him to see it now. For the internalist, however, there is no difficulty here. The epistemology of internalist reasons is rather straightforward: knowing what desires one has and knowing which of these would survive sound deliberation. There is no difficulty in explaining how an agent could come to know that he has some new desire, nor is there any difficulty in explaining *how* an agent comes to have a new desire.

My strategy was to devise an example where the agent is in a position to know in advance that he *will* have an internal reason, without being in a position to know that he will have the corresponding external reason. This allows us, so to speak, to subtract the internal reason from the external one. Given that the agent already knew he will have an internal reason, *whatever else* he later comes to know, then, is exactly the non-reductive answer to Williams's challenge.

We are now also in a position to reject these three internalist claims:

Before exposure to pain, we saw that internalism gives the wrong answers about what it would be rational for Zeno to do with respect to his future pain.

After exposure to pain, we can still distinguish internal and external reasons because they don't enjoy the same epistemic modality.

Finally, regarding *the point of the first exposure to pain*, internalists are in no position to complain that externalists cannot explain why Zeno couldn't see the reason earlier, and how he can see it now. In fact it is quite clear what happened: earlier he didn't know what pain feels like, and now he does.

Perhaps what Zeno comes to know, and how he comes to know it, will nevertheless still appear queer to some. After all it is a state that would be avoided "by anyone [rational] who was acquainted with it, not because of any contingent fact that this person, or every person, is so constituted that he desires this end, but just because the end has to-be-pursuedness somehow built into it."³³ All I hope to have shown is that reference to external reasons is

³³ Mackie 1977, 40. I've added the qualification that a *rational* agent would seek to avoid pain once acquainted with it, since Mackie unhelpfully describes the intrinsic *normative* force of objective value in motivational terms.

indispensable to making sense of the example of Zeno. If such reasons, and the state that gives them, are queer, then this is something we'll have to live with.

II.

There is a strong intuition that Zeno would acquire new normative knowledge when he first feels pain – when he first dislikes a present sensation of pain. And it seems clear enough that he would acquire this new normative knowledge by acquiring new *phenomenal* knowledge. Zeno comes to learn why a certain mental state is intrinsically bad: he comes to know that it's bad because it feels like *this*. All this suggests that *pain's badness resides in the way that it feels*.

It is clear how this could be so if the Sensation Theory were true: pain's badness would reside in the way a certain bodily sensation feels. But we saw that there are good reasons to reject the Sensation Theory. Indeed we are assuming that Zeno already knows what the sensation of pain feels like. So what, on the Dislike Theory, is left for Zeno to find out?

It is often assumed that the Dislike Theory implies the Pure Dislike Theory – the view that pain's badness is due to a conative attitude, not to its phenomenal character. But it does not follow, from the claim that pain is bad only when it involves affect or motivation, that the badness of the *state* of disliking the sensation of pain has nothing to do with what it's like to be in *this* state. It is a mistake to assume that to have an experience with a certain qualitative feel, and to have a sensation, are necessarily one and the same thing. To be sure, all sensations are also experiences. But many experiences are not sensations. Conscious thoughts are experiences, and so are the felt aspects of emotions and desires. I am suggesting that the overall state of disliking a present bodily sensation has phenomenal properties that are not exhausted by those contributed by the neutral sensation of pain. On the view I shall call the *Experiential Dislike Theory*, the state that is intrinsically bad is not that of having a sensation of pain, but that of suffering – of having this sensation and disliking it. And what makes *this* state bad is what it feels like. Disliking a sensation is a distinctive mental state, not merely the state of having a sensation as the object of a generic desire. This, I believe, is a better version of the Dislike Theory.

This view is supported by my example of Zeno. But it is also strongly supported by independent considerations. It is a very natural assumption that pain is bad because of what it feels like. This assumption is commonly denied because it is often assumed that it commits us to the discredited Sensation Theory. And although this assumption is commonly denied, the consequences of this denial are extremely counter-intuitive. Consider one obvious implication of the Pure Dislike Theory: the possibility of *hedonic inversion*, where someone would be in an identical overall state of consciousness we are in when we feel, say, an awful migraine or great ecstasy, yet for whom the value of that state is

reversed. It's very hard to believe that this is a genuine possibility.^{34,35}

OBJECTIONS

I.

The core of the Normative Knowledge Argument ran as follows. The first premise was the possibility of complete hedonic innocence. In its strongest form, it's the claim that

(1a) It is possible for a person to live all his life without ever experiencing a hedonic state

A less demanding but sufficient claim that I take to be implied³⁶ by (1a) is that

(1b) It is possible for a person to be in a state of (i) not feeling any hedonic state, (ii) nor being able to experientially remember or imagine what a hedonic state is like

I called the imaginary person who meets conditions (1a) or (1b) Zeno. It's uncontroversial that from (1b) it follows that

(2) Zeno wouldn't know what it's like to feel pain

Which, on the Dislike Theory, would amount to the claim that

(2b) Zeno wouldn't know what it's like to dislike a present sensation of pain

Zeno, however, knows a number of things about pain. In particular,

³⁴ But isn't masochism an actual example of hedonic inversion? Masochists claim to enjoy physical pain (at least in certain circumstances). This claim is ambiguous. If masochists enjoy feeling the *sensation* of pain, then they are not in the same state we are in when we suffer from pain. If they enjoy *disliking* this sensation, then they are taking pleasure in an experience that is in itself still intrinsically bad. Masochism is not an example of hedonic inversion on either reading.

³⁵ See Kahane (forthcoming) for further defense of this understanding of pain.

³⁶ Implied, that is, assuming nobody had implanted a memory of what pain feels like in this person's brain.

(3) Zeno knows that when he will be in pain, he will be in a state of disliking a present sensation

(4) If internalism is true, Zeno is already in a position to know all of pain's intrinsic normative properties

According to internalism, when Zeno will be in pain, he will have an internal reason to try to end it, and Zeno is already in a position to know that. The fact that Zeno doesn't know what it's like to dislike a sensation of pain is compatible with (4). Zeno already knows what the sensation of pain feels like and he knows he will dislike this sensation when he first feels pain (or will desire it to end after correct procedural deliberation). And he doesn't need to know what this further state feels like to know what internal reasons he will have when he first feels pain.

However, the Normative Knowledge intuition reveals that

(5) There is something Zeno doesn't know about pain's intrinsic normative properties, something he will come know only when he first feels pain

Zeno wouldn't know why we are justified in wanting to prevent, end or at least minimise pain, or what justifies a range of other attitudes and normative practices to do with pain. Another way of putting this is to say that Zeno wouldn't know what *makes true* the propositions (A) and (B).³⁷ The motivational facts cannot on their own explain why these propositions are true.

Therefore

(6) Internalism is false

Recall that I am using 'internalism' in the broadest sense. It is meant to cover any view that grounds practical reasons and value in the agent's present motivational dispositions. This includes a very wide range of views in meta-ethics: any view of practical reason or value that holds that normative

³⁷ Note that (A) and (B) are claims about agent-relative badness and reasons. It is obviously possible to know what pain is like yet deny that one has any moral reasons to care about and relieve the pain of others. The epistemology of these further, moral reasons would require a further explanation, even if, as Thomas Nagel claims, we could come to know them simply by reflecting on the nature of the experience (1989, 159-160)

knowledge can be derived a priori from knowledge about (i) a person's motivational states, and (ii) formal constraints or procedures governing practical reason. On all such views, the innocent Zeno should be in a position to know all of the normative facts. There would be no new normative knowledge for him to acquire when he first feels pain.³⁸

II.

We are now in a position to consider objections to the Normative Knowledge Argument. I take it that the most vulnerable premises are (1), (4) and (5). I'll now consider possible objections to each of these.

(1) Denying the possibility of Hedonic Innocence (premise 1)

It would be impossible for a rational agent to be free of hedonic experience if there was a necessary connection between states of valuation, affect and motivation and appropriate hedonic states.³⁹ There certainly is *a* connection. It is unpleasant when a strong urge or desire of ours is frustrated. Loss of something valuable can be painful. But what is the nature of this connection? It is doubtful that it is a necessary connection. What prevents us from conceiving of an agent who desires, deliberates, and values without ever feeling good or bad? And even if there is such a connection, and even if it's necessary, it seems to operate only at the level of *significant* frustration or *serious* disvaluing. We do not feel bad *whenever* a desire of ours is frustrated, or whenever we believe something bad has happened. And we can imagine that Zeno has lived an extremely sheltered life and that his desires are mild in strength and were never frustrated to any disturbing degree.

As I've noted earlier, the assumption of complete hedonic innocence is already a significant and non-obvious concession that I am making to the internalist, and I think I've said enough to make the possibility such innocence plausible. But there may still be those who nevertheless suspect that it is ultimately unintelligible to suppose that a rational agent could be utterly

³⁸ The argument can be cast to cover even more extensive ground. It has force against any view on which Zeno would be in a position to know the intrinsic evaluative and normative facts about pain on the basis of what he knows in the state of hedonic innocence – for example, views that try to explain pain's badness by reference to the fact that it is typically caused by bodily damage or that it represents such bodily damage.

³⁹ Hedonic innocence is incompatible with psychological hedonism, if that is understood as a claim about the motivation of all possible rational beings. And since I take Zeno to have full possession of evaluative and normative concepts, the claim is perhaps also incompatible with analytic normative hedonism and some empiricist accounts of the acquisition of evaluative or normative concepts. But these views, besides having nothing to do with the internalism/externalism debate, are hardly held by anyone.

ignorant of what suffering is like, and that consequently the force of any argument based on such a scenario is greatly weakened. So let me point out that the example of Zeno is only a dramatic device. The argument can be made to stand without it. Even an agent who has experienced pain can stand in essentially the same relation that Zeno stands to the first moment of pain. Even if the agent knows what it is like to suffer, he is still in a position to legitimately ask, of a complete description of a future state in which he will be strongly motivated to want a sensation to stop, whether that state is bad in any way. It seems to me that such an agent wouldn't be in a position to *know*, just from the kind of information available to Zeno, that this future state is identical with suffering. Of course, the intuitive force of the example is somewhat weakened because such an agent would certainly be in a position to *guess* that this state would be an unpleasant experience. But this doesn't make things any better for internalists, since from their point of view there is nothing left for this agent to guess.

(2) *Denying that internalism implies that Zeno is already in a position to know all of pain's intrinsic normative properties (premise 4)*

Internalists might be able to explain how Zeno gains new normative knowledge if they could show that Zeno, in his state of hedonic innocence, was missing a relevant piece of non-normative knowledge. On the internalist view, the normative facts are determined by what would motivate Zeno if he possessed full (non-normative) information and followed a sound deliberative route. If there is a non-normative fact about pain that Zeno doesn't know before first feeling pain, it *wouldn't* be clearly true that on their view Zeno is already in a position to know all of pain's intrinsic normative properties. The Normative Knowledge intuition would therefore be compatible with internalism.

And indeed internalists can point out to such a piece of knowledge: knowledge of what it's like to feel pain. Now this objection needs to take account of the distinction between the neutral sensation of pain and the experience of suffering. We saw that nothing prevents the hedonically innocent Zeno from knowing what the *sensation* of pain is like. But the Normative Knowledge Argument does rely on the fact that Zeno doesn't know what it's like to *dislike* this sensation. And if this is, as I've argued, itself a distinct form of experience, then here we have another piece of non-normative information that Zeno didn't know, allowing the internalist to run the objection.

There are several problems with this move. Notice first that in making it, the internalist would be giving up the Pure Dislike Theory. He would have to concede that disliking a bodily sensation is itself a distinct form of experience, and that it is *this* experience that is bad or reason-giving.⁴⁰ But if *this* experience

⁴⁰ The suggestion that there is a distinctive phenomenal feel to the state of disliking pain

is to be bad or reason-giving in a way that internalist can explain, then he must postulate yet another desire for *this* experience to end, a desire over and above the original dislike. This further desire, however, would have to be contingent; without it the state of disliking a bodily sensation wouldn't be bad, and so nothing prevents Zeno from knowing, in his innocent state, what *this* neutral state feels like, and we can again rerun the Normative Knowledge Argument. In any case, for the internalist to postulate such a further desire would be utterly *ad hoc*. There's no reason to suppose there is such a hidden second-order desire.

This objection suffers from an even more serious flaw. To be sure, according to internalism the reasons we have depend on what we would want after a sound deliberative route *and* full empirical information. And it may seem that Zeno in his innocent state is missing some empirical information. But this is irrelevant. Zeno *knows* that he will desire that experience to stop (or would desire this after deliberation). So he knows he will have an internal reason to try to stop it. Not knowing what exactly that experience would feel like, given that how this would feel like has no intrinsic normative significance, seems immaterial. So even if successful, this move doesn't really explain how Zeno could acquire new normative knowledge.

(3) *Denying that Zeno gains any normative knowledge when first exposed to pain (premise 5)*

It might be argued that the knowledge that Zeno acquires when he first experiences pain is not knowledge of a further normative fact, but a form of non-propositional knowledge.⁴¹ This of course needs to be non-propositional knowledge that is not normative (claiming acquaintance with a Moorean value property won't really help). On a natural version of this suggestion, the knowledge Zeno acquires is not knowledge *that* pain is bad, and that it's bad because it feels like *this*, but a form of *know-how*. Before he experienced pain, Zeno only had an intellectual grasp of the state he will be in when he finally feels pain. It is only by experiencing pain that he comes to vividly grasp it from the first-person. There is a difference between knowing that you will desire something and actually being in the felt grip of this desire. This, it might be argued, is the only divide that Zeno crosses. There are no external reasons that suddenly come into view when the metal needle first enters his body. There is only a change in practical orientation that has been misinterpreted as the acquisition of new normative knowledge.

To be sure, there may be a difference between knowing that one will have a desire and being in the felt grip of that very desire. But if a person who enters

but that this feel plays *no* part in making pain bad is so *ad hoc* and counterintuitive that it is hard to take seriously. But see the discussion in Kahane (forthcoming).

⁴¹ Such a move would parallel one common response to the Jackson's Knowledge Argument. See Nemirow, 1980 and Lewis, 1990.

such a state only acquires know-how, what is it exactly that he comes to know how to *do*? More importantly, there is nothing particularly unusual about knowing in advance that you will come to have a certain desire. Yet in such everyday cases there is no temptation to say that once we actually feel the desire we obtain any kind of new knowledge. It is hard to see why the case of pain should be so different.

META-ETHICS AND THE PROBLEM OF CONSCIOUSNESS

I.

The Normative Knowledge Argument is deliberately modeled after Frank Jackson's Knowledge Argument against physicalism. The dispute there is about phenomenal knowledge. The intuition is that Mary, an imaginary brain scientist who has complete knowledge of the neuroscience of colour vision but who grew up in a monochromatic room without ever having experienced colours, would come to learn something new when she first leaves the room and perceives, say, a red apple. Supporters of qualia claim that what Mary learns is a non-physical fact: what it feels like to experience the qualia of redness. Their physicalist opponents need to show that this gloss on our intuitions is mistaken. Either there is no knowledge gained, or the knowledge is non-propositional, or, despite appearances, it *is* knowledge of a physical fact. We do not need to go into the details.

There are two ways in which the Normative Knowledge Argument is less demanding than its predecessor. First, it is entirely clear what, according to the internalist conception, Zeno knows before he first feels pain. It is far less clear what it means to possess 'complete' knowledge of the physical and functional correlates of colour vision and what does, or does not, follow from such knowledge. Second, the internalist must cash out any change that Zeno undergoes when he first feels pain in motivational terms. But change of desire is common, whereas there is no equivalent range of familiar cases with which to contrast the physicalist's account of Mary's first encounter with redness.

We saw that it is natural to suppose that if Zeno acquires new normative knowledge when he first feels pain, he does so *by* acquiring new phenomenal knowledge. Does this mean that by claiming that Zeno acquires normative knowledge we are committing ourselves to a particular position in the debate about qualia? I would naturally prefer to be cautious in replying to this question. Presumably it commits us to a *cognitive* reading of that phenomenal knowledge claim, though I would hope to nothing more.

The main lesson I want to draw from the comparison between the Normative Knowledge Argument and Jackson's Knowledge Argument is not substantive but methodological. Think of what is supposed to be at stake in that

latter argument. One party holds that the sensation of redness or pain can be explained in terms of physical or functional states. The opposing party argues that such accounts leave out something vital: what redness or pain feel like. It would be silly to demand of the latter party to explain what it is exactly that is left out in terms that are themselves physical or functional. The only question is whether convincing examples can be set up where it seems clear that the physicalist story falls short of the facts. Externalists should take a similar line in reply to Williams's challenge. All externalists need to do is set up examples where it is clear enough that the internalist story leaves out something essential. In this paper I have tried to supply such an example.

II.

The Normative Knowledge Argument makes two main claims about Zeno's epistemic state before the first pain: that he doesn't know what pain feels like, and that he knows pain's motivational properties. The argument is thus compatible with physicalism since it leaves it open that Zeno *could* come to know what pain feels like, and thus whatever further normative properties this knowledge reveals, simply by reflection on the relevant *neuroscientific* facts. In other words, the Normative Knowledge Argument doesn't presuppose the soundness of Jackson's Knowledge Argument. At most, it suggests the falsity of some forms of functionalism.

It is worth, however, to briefly consider the prospects of a more ambitious variant of the Normative Knowledge Argument that *would* include a variant of the Knowledge Argument against physicalism. In this version of the argument, there is a quite a bit more that Zeno knows before the first pain. We now also suppose that Zeno is an expert in the *neuroscience* of pain. He knows not only pain's functional role but also its exact neural correlates – say, 10hz oscillations in the anterior cingulate cortex.⁴² Now the same intuitions that drive many to the view that Mary would gain new knowledge when she first sees something red should also drive many to a parallel view about Zeno and the phenomenal quality of suffering. But here the argument against physicalism would be taken one step further: it would also include the claim that at least some normative facts can't be known on the basis of physical knowledge, however complete.⁴³

⁴² This candidate neural correlate of pain was suggested by the neuroscientist John Stein in unpublished research on deep brain stimulation for the alleviation of neuropathic pain.

⁴³ Notice that the very fact that we can apply a form of the Knowledge Argument to the case of suffering is itself strong evidence that the experiential dimension of suffering plays an essential role in making it bad, and in saying this we need not commit ourselves to any particular metaphysics of experience. So we have further reason to reject what I called the Pure Dislike Theory. Even this weaker claim would not be congenial to many. It would be nicer if discussion of the normative dimension of pain could proceed without the intrusion of the perennial mind-

This form of argument has an ancient pedigree, although it's been long neglected. The earliest version of it I know of is in Sextus Empiricus, who wrote that

if ... the end consist in pleasure and ... the soul, like all else, is composed of atoms, it is impossible to explain how in a heap of atoms there can come about pleasure and assent or judgement that this object is choiceworthy and good, that object to be avoided and evil.⁴⁴

And Sidgwick wrote that

... so long as we confine our attention to their corporeal aspect – regarding them merely as complex movements of certain particles of organised matter – it seems impossible to attribute to these movements, considered in themselves, either goodness or badness. I cannot conceive it to be an ultimate end of rational action to secure that these complex movements should be of one kind rather than another, or that they should be continued for a longer rather than a shorter period. In short, if a certain quality of human Life is that which is ultimately desirable, it must belong to human Life regarded on its psychical side, or, briefly, Consciousness.⁴⁵

These passages can be understood as asserting a direct metaphysical claim: the claim that mere physical states couldn't possess the intrinsic value we ascribe to hedonic states. But they can also be read as making an epistemic claim: the claim that it is impossible to know that pain is intrinsically bad and reason-giving on the basis of exhaustive knowledge of the physical facts about pain. This wouldn't just be the claim that pain's intrinsic badness can't be *analytically derived* from these physical (or neural and functional) facts. There's nothing especially new or surprising about *this* kind of claim. Many have denied that empirical facts on their own imply evaluative or normative ones. The claim is rather that it is impossible to *know* the evaluative and normative facts about pain on the basis of these physical facts. No amount of reflection on patterns of neural activation would reveal that these patterns of activation are intrinsically bad. By contrast, simple exposure to the experience of pain is sufficient for knowledge of its badness.

Responding to a parallel claim, Mark Johnston remarks

If physicalism were true... all facts would just consist in facts about fundamental particles. Considered in themselves, these facts about particles would have no rational or moral importance. [Must we then] conclude that nothing has any importance? This is no a proof of nihilism. It is a *reductio ad absurdum*.⁴⁶

body problem. This would be the case if the Pure Dislike Theory were true. But it isn't.

⁴⁴ Sextus Empiricus 1933/2000, 252.

⁴⁵ Sidgwick 1907, 396.

⁴⁶ Johnston 1997. Johnston is criticising Parfit's claim that personal identity is

It might be replied that if pain is bad, and physicalism cannot account for this, then this is not proof of normative nihilism but of the falsity of physicalism. It is not, after all, utterly implausible to think that all intrinsic value supervenes on 'psychical' facts. This, however, is but a sketch of a possible argument against physicalism. My aim here was not to develop such an argument. One consequence of the Normative Knowledge Argument is that questions about the nature and epistemology of practical reason and value may turn out to be entangled with some of the deepest puzzles about the nature of consciousness. This sketch of an argument suggests that the dependency may also run in the opposite direction.^{47,48}

REFERENCES

- Altham, J. E. J. and Harrison, R. (eds.) 1995 *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge: Cambridge University Press.
- Brink, D. 1997. "Rational Egoism and the Separateness of Persons" in Dancy 1997.
- Campbell, J. 2002. *Reference and Consciousness*. Oxford: Oxford University Press.
- Chalmers, D. 1997. *The Conscious Mind*. Oxford: Oxford University Press.
- Clark, A. 2005. "Painfulness is not a Quale" in M. Aydede, (ed.) 2005. *Pain: New Essays on Its Nature and the Methodology of Its Study*. MIT Press.
- Dancy, J. (ed.) 1997 *Reading Parfit*. Oxford: Blackwell Publishing.
- Danziger, N., and Willer, J.-C. 2005. "Tension-type Headache as the Unique Pain Experience of a Patient with Congenital Insensitivity to Pain" . *Pain*, **117**, pp. 478-483.
- Gibbard, A. 2003. "Reasons Thin and Thick: A Possibility Proof" . *Journal of Philosophy*

unimportant.

⁴⁷ One of the few to hint at this possibility is Campbell 2002; 138. David Chalmers briefly considers the thought that moral properties (by which he seems to mean evaluative and normative properties quite generally) might raise the same problem for physicalism that he thinks is raised by phenomenal properties. But he goes on to remark that "moral facts are not phenomena that force themselves on us. When it comes to the crunch, we can deny that moral facts exist at all... The same strategy cannot be taken for phenomenal properties, whose existence is forced upon us." (Chalmers 1997, 83-84) This is an odd remark. The badness of pain seems to force itself upon us just like phenomenal properties. Indeed it imposes itself on us *through* a phenomenal property!

⁴⁸ I am very grateful to Nick Shackel, S. Matthew Liao, Timothy Chan and an anonymous referee for helpful comments and to Derek Parfit and John Broome for suggestions about an early draft.

100: 6, pp. 288–304.

Hare R. M. 1964. "Pain and Evil". reprinted in *Essays on the Moral Concepts*. London: Macmillan.

Hurshouse, R. Lawrence, G. and Quinn, W. (eds.) 1995. *Virtues and Reasons: Phillipa Foot and Moral Theory*, Oxford: Oxford University Press.

Jackson, F. 1982. "Epiphenomenal Qualia". *Philosophical Quarterly*. **32**, pp. 127–36.

Johnston, M. 1997. "Human Concerns without Superlative Selves" in Dancy 1997

Kahane, G. (forthcoming) "Pain, Dislike, and Experience". *Utilitas*.

Lewis, D. 1990. "What Experience Teaches" reprinted in Lycan 1990.

Ludlow, P., Nagasawa, Y. and Stoljar, D. (eds.) 2004. *There's Something About Mary*, MIT Press.

Lycan, W. G. (ed.) 1990. *Mind and Cognition*. Oxford: Basil Blackwell.

Macdonald, C. Smith, B. and Wright, C. (eds.) 1998. *Knowing Our Own Minds*. Oxford: Oxford University Press.

Mackie, J. L. 1977. *Ethics: Inventing Right and Wrong*. London: Penguin.

McDowell, J. 1995. "Might There be External Reasons?" in Altham and 1995.

Nagasako, E. M., Oaklander, A. L., Dworkin, R. A. 2003. "Congenital Insensitivity to Pain: An Update". *Pain*, 101 213-219.

Nagel, T. 1989. *The View From Nowhere*. Oxford: Oxford University Press.

Nemirow, L. 1980. review of T. Nagel, *Mortal Questions*. *Philosophical Review*. **89**, pp. 475 – 6.

Parfit, D. 1997. "Reasons and Motivation". *Proceedings of the Aristotelian Society*, Supplementary Volume **71**, pp. 99-130.

Quinn, W. 1995. "Putting Rationality in Its Place". in Hurshouse, Lawrence, and Quinn 1995.

Scanlon, T. M. 1998. *What We Owe To Each Other*. Cambridge, Mass.: Harvard University Press.

Sextus Empiricus, 1933/2000. *Outlines of Pyrrhonism* (trans. R. G. Bury). Cambridge, Mass.: Harvard University Press.

Sidgwick, H. 1907 *The Methods of Ethics* (7th ed.). London: Macmillan.

Williams, B. 1981. "Internal and External Reasons" in *Moral Luck*. Cambridge: Cambridge University Press.

----. 1989. "Internal Reasons and the Obscurity of Blame" reprinted in Williams, B. 1995. *Making Sense of Humanity*. Cambridge University Press.

----. 1995. "Replies" in Altham and Harrison 1995.