Philosophia (2006) 34:173–188
DOI 10.1007/s11406-006-9023-y

Modal Stability and Warrant

Colin P. Ruloff

Published online: 15 November 2006

Abstract Keith DeRose believes that it is a strength of his contextualist analysis that it explains why the recently much-discussed skeptical Argument from Ignorance (AI) is so persuasive. Not only *that*, however; DeRose also believes that he is able to explain the underlying dynamics of AI by utilizing *solely* the epistemological and linguistic resources contained within his contextualist analysis. DeRose believes, in other words, that his contextualist analysis functions as a genuinely *self-contained* explanation of skepticism. But does it? In this paper I argue that DeRose's analysis does *not* function as a self-contained explanation of skepticism since, as it turns out, DeRose's analysis is simply *irrelevant* to the main concerns of the skeptic. To the extent that DeRose's analysis is irrelevant in this way, I conclude that such an analysis cannot be considered a satisfactory treatment of AI.

Keywords warrant · skepticism · internalism

The Puzzle Generated by AI

A number of epistemologists have recently discussed some version of the following hypothesis:

> H: I am a disembodied brain floating in a vat of nutrient fluid being electrochemically stimulated to have the very perceptual experiences that I am currently having.

C. P. Ruloff (✉)
Department of Philosophy, Kwantlen University College,
12666 72nd Avenue, Surrey, British Columbia V3W.2M8, Canada
e-mail: colin.ruloff@kwantlen.ca

C. P. Ruloff
University of North Texas, TX, USA

Say that I am a "Brain in a Vat" or, for short, a "BIV", if and only if hypothesis H is true. Now consider the following skeptical Argument from Ignorance (AI), as DeRose (1995) calls it:

The Argument from Ignorance (AI)

(S_1) I do not know that I am not a BIV.
(S_2) If I do not know that I am not a BIV, then I do not know that I have two hands.
(S_3) Thus, I do not know that I have two hands.

Now, AI generates a puzzle. AI generates a puzzle since (S_1) and (S_2), which are both at least initially plausible, entail a conclusion whose denial, not-(S_3), we *also* find at least initially plausible. (S_1) is at least initially plausible because it seems that in order to know that I am not a BIV, I must *eliminate* or *rule out* the possibility that I am a BIV. But, given H, the BIV and I have perceptual experiences that are *phenomenologically indistinguishable*; It seems to the BIV, just as it seems to me, that he has two hands, that he is sitting at his desk, that he is typing away at his computer, and so on. Thus, given my current perceptual experiences, I *can not* eliminate the possibility that I am a BIV. But then, it seems, I do not *know* that I am not a BIV. These sorts of considerations contribute to (S_1)'s plausibility. But (S_2) is also at least initially plausible because it seems that if I can not eliminate the possibility that I am a BIV – that is, if it really is an *open question* as to whether or not I am a BIV – then I can not properly claim to know that I *do* in fact have two hands. But at the same time, it also seems that (S_3)'s denial is true and that I *do* know that I have two hands. So two plausible premises – (S_1) and (S_2) – entail a conclusion whose denial we also find intuitively plausible.

AI, then, appears to generate a puzzle of sorts; (S_1), (S_2), and not-(S_3) are independently plausible, but, it seems, jointly inconsistent. It would seem, therefore, that one of (S_1), (S_2), and not-(S_3) has got to be false. But which one? Whereas the skeptic will happily endorse (S_3), the non-skeptics tend to divide themselves into two main camps; either, following Moore, they deny (S_1), or, following Nozick, they deny (S_2). In contrast, the currently fashionable *contextualist* analysis – as advanced by Cohen (1986), DeRose (1995), Lewis (1996), and others – claims to offer a solution to the puzzle generated by AI which not only preserves the intuition that each of (S_1), (S_2), and not-(S_3) is true, but also explains *why* (S_1), (S_2), and not-(S_3) need not necessarily result in a puzzle in the first place.

So just what *is* the contextualist analysis of AI? Stated generally, contextualists maintain that the truth-values of knowledge-ascriptions are in some important sense *context-dependent*. More precisely, according to the contextualist, certain features of a context - features such as the communicative intentions of the members of a conversational context – somehow *fix* or *determine* the standards that one must meet in order for one's beliefs to count as knowledge (Black, 2006). This allows for the possibility that different conversational contexts fix different epistemic standards. In some conversational contexts – specifically, contexts in which we are seriously entertaining radical skeptical hypotheses such as the BIV hypothesis – the epistemic standards are unusually high, and it is exceedingly difficult for our beliefs to count as knowledge in such contexts. In these so-called *high-standards* contexts, AI presents us with a sound argument. In most conversational contexts, however, the epistemic standards are comparatively quite low, and our beliefs can and often do count as knowledge in these contexts. In these so-called *low-standards* contexts, we can safely reject AI since its first premise, (S_1), will be false. So, on the contextualist analysis, in high-standards contexts AI presents us with a perfectly sound argument, but in low-standards contexts, this same

argument, although deductively valid, is *un*sound for its first premise will be false. But notice: Given that there is no single context in which (S_1), (S_2), and not-(S_3) all semantically express true propositions, the contextualist analysis keeps the conflict between (S_1), (S_2), and not-(S_3) from presenting itself. Accordingly, say contextualists, the puzzle generated by AI vanishes.

Since Keith DeRose's view is perhaps the most fully developed contextualist solution to the puzzle generated by AI, this will be the view that I discuss in what follows.

The Subjunctive Conditionals Explanation of (S_1) of AI

DeRose believes that in order to properly solve the puzzle generated by AI we need to first provide an explanation as to why we are inclined to regard AI's first premise as plausible. DeRose explains the intuitive plausibility of (S_1) of AI by appealing to Robert Nozick's (1987) subjunctive conditionals account of knowledge. According to that account:

> we have a very strong general inclination...to think that we don't know that *p* when we think that our belief that *p* is a belief we would hold even if *p* were false. Let's say that *S*'s belief that *p* is *insensitive* if *S* would believe that *p* even if *p* were false....We tend to judge that *S* doesn't know *p* when we think that *S*'s belief that *p* is insensitive. (DeRose, 1995, p. 18)

In other words, according to DeRose, we tend to judge that *S* does not know that *p* when we think that *S* would believe that *p* even if *p* were false (so that *S*'s belief is "insensitive"). Expressed in terms of possible worlds, DeRose's basic idea is that we tend to judge that *S* does not know that *p* if *S* believes *p* in any of the nearest not-*p* worlds. For our purposes, we can capture DeRose's general point somewhat more fully and formally by way of the following modal stability constraint:

> [Modal stability constraint (MSC)] An agent *S* knows a proposition *p* if and only if *S* believes *p*, *p* is true and *S*'s belief that *p* is *modally stable* – i.e., *S* believes *p* in those worlds centered around the actual world in which *p* is true, and *S* does not believe that *p* in the nearest world centered around the actual world in which *p* is false.

In order for *S*'s belief that *p* to count as an instance of knowledge on MSC, *S*'s belief that *p* must be *modally stable* – that is, *S*'s belief that *p* must not only be true in the actual world and in those worlds surrounding the actual world, but, crucially, *S*'s belief that *p* must also "track the truth" in such a way that in the nearest world in which *p* is false, *S* does not believe *p*. If *S*'s belief that *p* fails to track the truth in this way, *p* is said to be modally *un*stable for *S*. Generally speaking, then, we might say that, on MSC, *S*'s belief that *p* counts as an instance of knowledge provided that *S*'s belief that *p covaries* with the fact that *p* in all worlds surrounding and including the actual world out to the nearest not-*p* world. Or to use DeRose's terminology, we might say that *S*'s belief that *p* counts as an instance of knowledge provided that *S*'s belief *matches the fact of the matter* in the sphere of worlds that ranges from the actual world out to the nearest not-*p* world.

DeRose (1995) believes that MSC offers a "powerful" (p. 19) and an "intuitively appealing" (p. 24) explanation as to why we are inclined to regard the first premise of AI as true. According to DeRose:

> the problem with my belief that I am not a BIV...is that I would have this belief (that I'm not a BIV) even if it were false (even if I were one). It is this that makes it hard to

> claim to *know* that I'm not a BIV. For, according to (MSC), we have a very strong general, though not exceptionless, inclination to think that we don't know that *p* when we think that our belief that *p* is a belief that we would hold even if *p* were false. (p. 18)

In other words, the reason as to why we are inclined to regard (S_1) of AI is true is that your belief that you are not a BIV is not modally stable; You would believe you were not a BIV even if you *were* a BIV; Although your belief that you are not a BIV is true in the actual world and in nearby worlds, it is difficult to claim to *know* that you are not a BIV, says DeRose, because in the nearest world in which this belief is false – i.e., the nearest world in which you *are* a BIV – you would *continue* to hold the belief that you are not a BIV. Again, using DeRose's terminology, although your belief that you are not a BIV does a good job of matching the fact of the matter in those worlds centered around and including the actual world, the reason as to why you do not *know* that you are not a BIV is that in the nearest world in which you are a BIV, your belief that you are not a BIV *fails* to match the fact of the matter.

Incorporating MSC

According to DeRose, then, MSC offers an elegant explanation of why we are inclined to regard (S_1) of AI as true. DeRose goes on to incorporate MSC into what is perhaps the most fully developed contextualist solution to the puzzle generated by AI. Schematically, and glossing over a number of details, DeRose's solution goes roughly as follows.

DeRose claims that when I assert "I know that I have two hands" in a conversational context in which no skeptical hypothesis are being seriously entertained, the epistemic standards that are in effect are quite weak and merely require, following MSC, that my belief that I have two hands track the truth from the actual world out to the nearest world in which I fail to have hands. The nearest world in which I am handless will serve as the *modal boundary*, as it were, for this context since it fixes the sphere of worlds across which my belief that I have two hands must track in order for it to be properly said that I know that have two hands. But this, according to DeRose, is a standard that I can readily meet; Since I *do* believe that I have two hands in the actual world and in nearby worlds, and *do not* believe that I have two hands in the nearest world in which I am handless, I *know*, according to DeRose, in this context that I have two hands; My belief successfully tracks the truth across the relevant sphere of worlds and is, accordingly, modally stable. Consequently, in this context, I speak truly when I assert "I know that I have two hands". But, continues DeRose, *in this very same context*, I also know that I am not a BIV; In the sphere of worlds relevant to this context – i.e., the sphere that extends from the actual world out to the nearest world in which I am handless – I both am not a BIV and *believe* that I am not a BIV. So, although my belief that I am not a BIV is itself not modally stable, this belief nevertheless has a sufficiently strong degree of *modal entrenchment*, as it were, in *this* context for me to know that I am not a BIV. Hence, in an everyday conversational context, the epistemic standards are such that not only do I know that I have two hands, but I *also* know that I am not a BIV.

According to DeRose, however, things are altogether different in conversational contexts in which the skeptic makes salient skeptical hypotheses, such as the BIV hypothesis. When

I assert "I know that I'm not a BIV" in a context in which the skeptic raises such a hypothesis, by the following rule of sensitivity (DeRose, 1995, p. 36):

> RULE OF SENSITIVITY: When it's asserted that *S* knows (or doesn't know) that *p*, then, if necessary, enlarge the sphere of relevant worlds so that it at least includes the closest world in which *p* is false.

the epistemic standards are driven upwards and the sphere of relevant worlds expands and now reaches out to those modally distant BIV worlds. In order for me to know that I am not a BIV in this high-standards context, my belief that I am not a BIV must now – again, following MSC – track the truth across an *expanded sphere* of worlds, namely, the sphere that ranges from the actual world out to the nearest world in which I *am* a BIV. The nearest world in which I am a BIV now serves as the modal boundary for *this* context. But, says DeRose, this is a standard that I unfortunately *can not* meet; Although my belief that I am not a BIV matches the fact of the matter in the actual world and in nearby worlds, I do not *know* in this context that I am not a BIV because in the nearest world in which I am a BIV, I would *continue* to hold the belief that I am not a BIV. Thus, since my belief fails to track the truth in this high-standards context (and is, accordingly, modally *un*stable), I speak falsely when I assert "I know that I am not a BIV." But, continues DeRose, in the sphere of worlds relevant to this conversational context, I *also* fail to know that I have two hands; Although my belief that I have two hands matches the fact of the matter in the actual world and in nearby worlds, I do not know that I have two hands because in the nearest BIV world in which I am both handless and bodiless I would *continue* to believe that I have two hands. More precisely, although my belief is modally stable in everyday conversational contexts, in contexts in which skeptical scenarios are salient, my belief that I have two hands does not have a *sufficiently strong* degree of modal entrenchment in order for me to know that I have two hands. In contexts in which the skeptic succeeds in raising the epistemic standards for knowledge – in contexts in which the sphere of worlds is unusually large – then, not only do I fail to know that I am not a BIV, but I also fail to know that I have two hands.

So how does all of this apply to the puzzle generated by AI? Roughly as follows: Since, by making salient the BIV hypothesis, the skeptic forces me into a context in which my belief that I am not a BIV is modally unstable, I do not, according to DeRose, know that I am not a BIV in such a context, i.e., the proposition expressed by (S_1) is true in such a context. But if, according to (S_1), I do not know that I am not a BIV, then, by modus ponens on (S_2) and keeping the sphere of worlds fixed, I also do not know in such a context that I have two hands, i.e., the proposition expressed by (S_3) is true in such a context. In high-standards contexts, then, the skeptic presents us with a sound argument. According to DeRose, however, in low-standards contexts, I do know such things as that I have two hands because this belief *is* modally stable. So, in low-standards contexts, the proposition expressed by not-(S_3) is true. But if not-(S_3) is true, then I can infer from not-(S_3) together with (S_2) by modus tollens that, contrary to the proposition expressed by (S_1), I know that I am not a BIV after all. Thus, in low-standards contexts, AI, although deductively valid, is unsound for its first premise will be false.

But now, if we are never actually faced with a direct conflict between (S_1), (S_2), and not-(S_3), why does it seem as if we *are* in fact faced with such a conflict? The general contextualist answer is this: Since we typically find ourselves in low-standards contexts, we naturally tend to evaluate knowledge-attributions according to the epistemic standards that are appropriate for those contexts. Thus, we tend to regard not-(S_3) as obviously true.

However, since (S_1) makes explicit reference to BIV's, our evaluation of (S_1) tends to lead us to entertain the exotic BIV hypothesis. Doing so raises the epistemic standards for knowledge – that is, it pushes us into a context in which the epistemic standards are unusually high (Black, 2006) – and so we tend to regard (S_1) as true too. And so it seems that we are faced with a direct conflict between (S_1), (S_2) and not-(S_3). Yet this conflict is merely apparent. For, as we have just seen, when the standards are low, (S_1) is false, while (S_3)'s denial is true. But, when the skeptic sneaks in a change of context by making salient the BIV hypotheses, the standards rise accordingly, and in this new context not-(S_3) is false, while (S_1) is true. The contextualist story, then, is that otherwise competent speakers "get fooled" (Bach, 2004, p. 17) or bamboozled because they do not notice when the epistemic standards get raised by the skeptic.

The Argument Initially Stated

As we will see in somewhat more detail below, DeRose believes that it is a strength of his contextualist analysis that it explains why the skeptic's AI is so persuasive. Not only *that*, however; DeRose also believes that he is able to explain the underlying dynamics of the skeptic's AI by utilizing *solely* the epistemological and linguistic resources contained within his contextualist analysis – by way of, for instance, the rule of sensitivity, the modal stability constraint, and so on. DeRose believes, in other words, that his contextualist analysis functions as a genuinely *self-contained* explanation of skepticism. But does it? In what follows I want to argue that DeRose's analysis does *not* function as a self-contained explanation of skepticism since, as it turns out, DeRose's analysis is simply *irrelevant* to the main concerns of the skeptic. To the extent that DeRose's analysis is irrelevant in this way, I conclude that DeRose's analysis cannot be considered a satisfactory solution to the puzzle generated by AI.

Some Core Features Isolated

Let us isolate a number of core features of DeRose's contextualist analysis.

Firstly, skeptical arguments such as AI seem to have a pull on our assent that cries out for explanation. One important advantage claimed by DeRose to favor his contextualist analysis is this; contextualism offers us a good explanation of *why* skeptical arguments such as AI are so persuasive. Contextualism, in other words, offers an explanation which captures the *underlying dynamics* of AI. In particular, DeRose believes that his contextualist analysis offers us a good explanation as to why the crucial first premise of AI – (S_1) – is true. Consider again DeRose's explanation. On DeRose's analysis, when the skeptic makes salient the BIV hypothesis, the rule of sensitivity is invoked, and the epistemic standards are driven up to such a level that my belief that I am not a BIV must be modally stable if this belief is to count as an instance of knowledge. But, says DeRose, in this high-standards context – in what Bach (2004, p. 14) calls a "skepistemic" context – my belief that I am not a BIV is not, and *could* not, be modally stable; I would believe that I am not a BIV even if I *were* a BIV. Hence, according to DeRose, (S_1) is manifestly true; I do not *know* that I am not a BIV. Let us say, then, that:

P1: On DeRose's contextualist analysis, what accounts for the truth of (S_1) of AI is that, by making salient the BIV hypothesis, the skeptic – by way of the RULE OF

SENSITIVITY – raises the standards for knowledge to such a level that my belief that I am not a BIV fails to be modally stable.

On DeRose's analysis, then, it is the fact that the skeptic places me in a high-standards context in which my belief that I am not a BIV *fails* to be modally stable which accounts for the truth of (S_1).

For our purposes, what is important to notice here is that, on DeRose's analysis, the skeptic is viewed as someone who, in presenting the BIV hypothesis, *raises the semantic standards for knowledge*. Of course,[1] the mere *presenting* of the BIV hypothesis by the skeptic need not automatically raise the standards for knowledge; it merely creates "some pressure" (Kornblith, 2000, p. 28) in that direction, to which the otherwise non-skeptic may or may not agree. But, says DeRose, in cases where the conversational partners are generally cooperative and go with the conversational flow, as it were, the skeptic *does* oftentimes succeed in raising the standards; In such contexts, says DeRose, the skeptic not only threatens to raise the standards for knowledge, but also *succeeds* in doing so. As DeRose (1992) notes, on his contextualist analysis, when the skeptic presents the BIV hypothesis, she oftentimes "manipulates various conversational mechanisms that *raise the semantic standards for knowledge*" (p. 917).[2] And again, DeRose (2000) claims that (S_1) of AI is persuasive precisely because the presentation of the BIV hypothesis by the skeptic has a strong tendency to "*put into play the very "absolute" standards*" (p. 134) at which I fail to know much if anything about the external world.[3] DeRose's thus accounts for our tendency to be seduced by skeptical argument through "an appeal to the pragmatics and dynamics of conversation" (Kornblith, 2000, p. 28). The central point to be kept in mind here, then, is this; On DeRose's analysis, the skeptic is someone who, in presenting the BIV hypothesis, oftentimes *does* succeed in raising the semantic standards for knowledge to a level at which (S_1) is true.

But now, says DeRose, if (S_1) of AI is indeed true in a high-standards context – and if we hold the sphere of worlds across which my belief must track fixed – then, by *modus ponens* on (S_2), I also do not know in such a context that I have two hands, i.e., (S_3) is true in such a context. Says DeRose (1995), "once the standards have been so raised, we correctly sense that we only could *falsely* claim to know such things as that we have hands" (p. 5).[4]

The fact that the skeptic is able to raise the standards in this way, however, does not mean that I can never correctly be said to have ordinary propositional knowledge. As we have already seen, DeRose wants to avoid the boldly skeptical conclusion that I do not know that I have two hands, and he does this by maintaining that in contexts in which the epistemic standards are comparatively low and in which no skeptical hypotheses are being seriously considered, I *do* know that I have two hands; I know that I have two hands because my belief that I have two hands *is* modally stable in such contexts. Thus, in low-standards contexts not-(S_3) is true. Let us say, then, that:

P2: On DeRose's contextualist analysis, what accounts for the truth of not-(S_3) of AI is that, in low-standards contexts in which no skeptical hypotheses are being seriously entertained, my belief that I have two hands *is* modally stable.

[1] As Kornblith (2000) has noted.

[2] Italics added.

[3] Italics added. And elsewhere, DeRose (1995) says that on his analysis, we should endorse (S_1) as true not at all standards, "but only at the *unusually inflated standards* conducive to skepticism" (p. 39).

[4] Italics added.

As DeRose (1995) states, the mere fact that the skeptic can invoke very high standards that we do not live up to has no tendency to show that we do not "satisfy the more *relaxed standards* that are in place in more ordinary conversations and debates" (p. 5).[5] But now, if not-(S_3) is true in such contexts – and, again, if we hold the sphere of worlds across which my belief must track fixed – then I can infer from not-(S_3) together with (S_2) that (S_1) is false in such contexts, i.e., I can infer that I *do* know that I am not a BIV.

On DeRose's analysis, then, the skeptic is viewed as someone who deploys what we might call the *modal forcing strategy*; In low-standards contexts we know quite a few things about the external world. We know such things because, generally speaking, our beliefs successfully track the truth across the sphere of worlds relevant to these kinds of conversational contexts. In high-standards contexts, however, we fail to know much of anything. By making salient well-formed skeptical hypotheses (such as the BIV hypothesis), the skeptic essentially *forces* the sphere of worlds to expand in such a way that in this context our beliefs *fail* to track the truth in the appropriate way so as to count as knowledge. Accordingly, it would not be inaccurate to call DeRose's analysis a *contextualized tracking analysis* of knowledge-attributions.

Secondly, it is important to bear in mind that not only is DeRose's contextualist analysis supposed to explain the underlying dynamics of AI, but DeRose's analysis is supposed to do so by utilizing *solely* the epistemological and linguistic machinery contained within his analysis – by utilizing *solely* the rule of sensitivity, the modal stability constraint, and so on. It is, then, natural to interpret DeRose as offering what he considers to be a genuinely *self-contained* explanation of skepticism, an explanation that does not appeal to any *non*-contextualist or *extra*-contextualist resources. As Johnsen (2001) has put the point, DeRose is aiming to explain the underlying dynamics of skepticism "*solely on contextualist grounds*" (p. 386).[6] If this is right, then DeRose appears to affirm the following self-containment thesis:

> **P3**: SELF-CONTAINMENT THESIS: DeRose's contextualist analysis is intended to serve as a *self-contained explanation* of the underlying dynamics AI, an explanation that does not appeal to any *extra*-contextualist or *non*-contextualist resources.

That DeRose's analysis is committed to something like P3 is not something that is always made very clear in the literature. It is worth stressing, however, that DeRose's version of contextualism is to be viewed in part as an *explanatory theory* – a theory that not only attempts to diagnose and solve the puzzle generated by AI, but one that also attempts to *explain* the persuasiveness of AI and of skepticism more generally.

The self-containment thesis appears to be crucial to the overall success of DeRose's contextualist analysis for, if DeRose's analysis *fails* to explain the underlying dynamics of AI by way of the epistemological and linguistic machinery contained within his analysis, then such an analysis can scarcely be called a "*solution*" to the puzzle generated by AI, as DeRose claims it to be.[7]

Thirdly, it is important to note that DeRose's contextualist analysis of AI is embedded within a broadly *externalist* framework. As we have seen, on DeRose's analysis, an agent *S* knows that *p* in a conversational context *c* provided that, roughly, *S* believes *p*, *p* is true, and *S*'s belief that *p* meets the standard for truth-tracking that governs *c*. On DeRose's

[5] Italics added.

[6] Italics added.

[7] See DeRose, 1992, p. 917, 1995, p. 5.

account, then, *S*'s belief that *p* counts as an instance of knowledge in *c*, *not* because *S* possesses any evidence or argument for *p*, *not* because *S* has introspective or reflective access to some state, property, or condition which somehow evidentially supports *p*, not because *S* is aware of any truth-indicating grounds for *p*, but *solely* because *S* holds *p* in certain possible worlds and not in others, *solely* because *S*'s belief that *p* matches the fact of the matter across the sphere of worlds relevant to *c*. DeRose is fairly explicit about his contextualist analysis being broadly externalist in character. When confronted with the question "*How* do we know that we're not BIV's?" DeRose (2000) answers:

> *not* by means of any argument, nor even by means of having any effective evidence to that effect. More positively, in my own case, my account is that we know...that we're not BIV's because our belief as to whether we're BIV's matches the fact of the matter in the actual world and in the sufficiently nearby worlds. (p. 135)

On DeRose's analysis, then, *S* can know that *p* even though *S* is totally ignorant of the fact that her belief that *p does* meet the standard for truth-tracking that governs *c*, even though, in other words, *S* is not in any way *aware* that her belief that *p* actually matches the fact of the matter in the sphere of worlds relevant to *c*. DeRose's analysis carries with it the implication that *S* need not be intellectually responsible for personally overseeing that the right sort of counterfactual connection holds between her beliefs and the world in order to be correctly said to know. Either the appropriate counterfactual connection between *S*'s belief and the world obtains, or it does not obtain, but this is not a fact of which *S* needs to be reflectively *aware*. Since the externalist aspect of DeRose's contextualist analysis has been insufficiently noted in the literature, let us be explicit and say:

> **P4**: DeRose's contextualist analysis is embedded within a *broadly externalist framework*; On DeRose's analysis, *S*'s belief that *p* counts as an instance of knowledge in a conversational context *c* if and only if *S* believes *p*, *p* is true, and *S*'s belief that *p* tracks the truth across the sphere of worlds relevant to *c*. On this analysis, *S* can know that *p* even though *S* is not in any way *aware* that her belief that *p* is evidentially supported by some body of evidence or truth-indicating grounds.

The Skeptic's Basic Argument for (S_1)

Having isolated four core features of DeRose's contextualist analysis, let us now shift our attention and focus on (S_1) of AI.

It is interesting to note that contextualists rarely, if ever, attempt to formally reconstruct the skeptic's argument for AI's first premise, (S_1). As Byrne (2004) has pointed out, contextualists often claim (or at least imply) that no argument is needed for (S_1), or if an argument is required, it merely consists in drawing attention to the BIV hypothesis, and then noticing the skeptical consequences that follow from entertaining such a hypothesis. In fact, according to many contextualists, when one contemplates (S_1) it becomes "pretty much *obvious*" (Byrne, 2004, p. 304) that one does not know that one is not a BIV. And this all seems to be true of DeRose as well; DeRose nowhere attempts to reconstruct the skeptic's basic argument for (S_1), implying that a formal reconstruction of the skeptic's argument for (S_1), whatever that argument may be, is somehow *unnecessary*.

Here I must assert that, although (S_1) of AI is at least initially plausible, I believe that (contrary to DeRose) the skeptic *does* indeed have a particular kind of argument for (S_1) in

mind – an argument which, as we shall see, is motivated by straightforwardly *internalist* considerations.

So, just what *is* the skeptic's argument for (S_1)? Which premises does the skeptic appeal to in order to arrive *at* (S_1)? Following closely the lead of Johnsen (2001), I believe that the skeptic's basic argument for (S_1) proceeds roughly along the following lines. Let "warrant" designate that property which, when conjoined with true belief, yields knowledge. With this in mind, it is plausible to see the skeptic as beginning with the following assumption: What an agent *S* has to go on in determining the nature of the external world consists solely of *S*'s *current conscious experience*. Or, to be more technically accurate, the skeptic holds that in order for *S*'s belief that *p* to be warranted – where *p* is a proposition regarding the nature of the external world – *S* must be in some important sense *aware* of that which supports the belief that *p*, *S* must be able to evidentially support *p* solely by way of what is *cognitively or introspectively accessible* to her. As Luper (2003) puts it, the traditional project of skepticism is motivated by the idea that we are warranted in believing that *p* provided that "we have *introspective access* to evidence that makes *p* more likely than any of the alternatives to *p*" (p. 15). Or, as Pritchard (2002) says, the skeptic is ultimately guided by the idea that "the positive epistemic status of an agent's belief is wholly determined by facts which she is in position to know by *reflection alone*" (p. 6). Indeed, the idea that warrant-conferring properties must be introspectively accessible properties is what Byrne (2004) has recently called a "platitudinous assumption" (p. 306) of the skeptic's reasoning for (S_1). On the traditional skeptical project, then, the evidence available to *S* regarding the nature of reality would consist of *S*'s current set of beliefs, perceptions, memories, sensations, ideas, impressions, or other mental entities generally accessible by way of introspection or cognitive reflection. Accordingly, it is plausible to see the skeptic as affirming what I will call the evidence-allowability constraint:

> EVIDENCE-ALLOWABILITY CONSTRAINT (EAC): An agent *S*'s beliefs about the nature of the external world are warranted only if *S* is able to evidentially support her beliefs by way of her current conscious experience, by way of what's accessible to *S* by way of introspection or cognitive reflection.[8]

Armed with the evidence-allowability constraint, the skeptic then goes on to point out that a wide range of hypotheses, including the BIV hypothesis, about the nature of the external world are perfectly *compatible* with *S*'s current conscious experience in that nothing in *S*'s current conscious experience evidentially supports any one of these hypotheses more strongly than any of the others. That is, the skeptic points out that that a wide range of hypotheses, including the BIV hypothesis, about the nature of reality are *consistent* with *S*'s current conscious experience in that nothing in that experience enables her to discriminate the state of affairs in which she is, say, a hapless BIV from the state of affairs in which she is at home typing away at her computer safely out of the reach of evil cognitive scientists. To use Lehrer's (2000) terminology, the skeptic notes that there is no *subjectively detectable difference* between the situation in which *S* is a bodiless and handless BIV and the situation in which is *S* not. In addition to the evidence-allowability constraint, then, the skeptic appears to affirm the following evidence-neutrality constraint:

> EVIDENCE-NEUTRALITY CONSTRAINT (ENC): A wide range of hypotheses, including the BIV hypothesis, about the nature of the external physical world are perfectly

[8] For more on the *Cartesian* underpinnings of AI, see Rowlands (2003), pp. 25–29.

consistent with *S*'s current conscious experience in that nothing in this experience evidentially supports or favors any one of these hypotheses over any of the others.

But now, says the skeptic, if nothing in *S*'s current conscious experience evidentially supports any one of these hypotheses, including the BIV hypothesis, more strongly any of the others – if, in other words, *S*'s current conscious experiences evidentially supports each of these hypotheses, including the BIV hypothesis, *to the same degree* – then *S*'s current conscious experience can not very well serve as *good* evidence for her belief that any particular one of these hypotheses is true or that any particular one of these hypotheses is false. In particular, if *S*'s current conscious experience supports each of these hypotheses, including the BIV hypotheses, to the same degree, then *S*'s current conscious experience can not serve as good evidence for *S* to back up her natural tendency to *reject*, for instance, the BIV hypothesis. As Greco (2000), Sosa (2003), Hill (1996), and Vahid (2005) have pointed out, the skeptic's reasoning at this point in the argument is motivated by an implicit appeal to the following adequate evidence constraint:

ADEQUATE EVIDENCE CONSTRAINT (AEC): For any agent *S* and empirical hypotheses h_1 and h_2, if *S*'s available evidence supports h_1 and h_2 to the same degree, then *S*'s available evidence does not serve as *good* evidence for *S* to prefer h_1 over h_2.

Since *S*'s available evidence, i.e., her current conscious experience, supports the BIV Hypothesis and what we might call the "Real-World Hypothesis" to the same degree, *S*'s available evidence simply does not serve as good evidence for *S* to believe that the Real-World hypothesis is true and that the BIV Hypothesis is false. But now, if *S*'s available evidence does not serve as good evidence for her belief that the BIV hypothesis is false, then, says the skeptic, *S*'s belief that this hypothesis is false can not be a *warranted* belief. But given, by hypothesis, that warrant is a *necessary* condition for knowledge, it follows that *S* does not *know* that the BIV hypothesis is false. But to say *that* is just to say that (S_1) of AI is true. Hence, (S_1) of AI is true. Formally, then, the skeptic's basic argument for (S_1) appears to have the following basic structure:

The Skeptic's Basic Argument for (S_1) of AI

1. My beliefs about the nature of the external world are warranted only if I am able to evidentially support these beliefs by way of my current conscious experience ***c***. [evidence-allowability constraint (EAC)]
2. A range of hypotheses, including the BIV hypothesis, about the nature of the external world are perfectly *consistent* with ***c*** in that nothing in ***c*** evidentially supports or favors any one of these hypotheses over any of the others, i.e., ***c*** evidentially supports these hypotheses to the same degree. [evidence-neutrality constraint (ENC)]
3. If ***c*** supports each of these hypotheses, including the BIV hypothesis, to the same degree, then ***c*** does not serve as good evidence for my belief that the BIV hypothesis is false.
4. But if ***c*** does not serve as good evidence for my belief that the BIV hypothesis is false, then my belief that I am not a BIV is not a *warranted* belief.
5. But a necessary condition for knowing that I am not a BIV is that my belief that I am not a BIV *be* a warranted belief.
6. Therefore, I do not know that I am not a BIV, i.e., (S_1) of AI is true.

The evidence-allowability constraint and the evidence-neutrality constraint both appear to be important components of the skeptic's argument for (S_1) of AI. Indeed, it is

exceedingly difficult to see how the BIV hypothesis could be a live hypothesis *in the first place* unless something very much like these constraints are both true. Accordingly, it seems reasonable to hold that this argument – which also bears an important overall structural similarity to arguments recently forwarded by Sosa (2003), Greco (2000) and, to a lesser extent, Vahid (2005) – is the skeptic's basic argument for (S_1) of AI. At the very least, it appears that the evidence-allowability constraint and the evidence-neutrality constraint are both *necessary* components of the skeptic's argument for (S_1).

The Skeptic as an Internalist

Let us briefly have a closer look at the evidence-allowability constraint. Notice that this constraint demands that, in order for an agent to be warranted in believing that *p*, the agent must be able to evidentially support *p* by way of what is introspectively or cognitively accessible to her. But notice: To say *this* is just to say that *internalism about warrant is true*; Internalism *just is* the claim that the properties that confer warrant on a belief must be "internal" to the mind of the agent who holds that belief – that the warrant-conferring properties of a belief must be properties which an agent can access or pick out just by way of cognitive reflection or introspection alone. Applied specifically to the BIV hypothesis, internalism demands that, whatever the nature of external world, the properties that confer warrant on *S*'s belief that she is not a BIV must be properties to which *S* has introspective or cognitive access. The problem, however, is that the BIV hypothesis is formulated in such a way that, from *S*'s cognitive perspective, there will be no subjectively detectable difference between the state of affairs in which *S* is a BIV and the state of affairs in which she is, say, sitting at home in front of her computer. Accordingly, there will be no properties introspectively or cognitively accessible to *S* that would indicate to her that she is *not* the victim of such a hypothesis. But then, *S* is not *warranted* in believing that she is not a BIV. But if *S* is not warranted in believing that she is not a BIV, then *S* does not *know* that she is not a BIV, i.e., (S_1) of AI is true. (Or so argues the skeptic.) So the skeptic is an *internalist* with respect to warrant. More exactly, it appears that it is a pretheoretic commitment to internalism that is ultimately motivating the skeptic's basic argument for (S_1). Pritchard (2002) makes this just this point. He states that a commitment to internalism will be "sufficient to motivate the skeptical premise [(S_1)] ... no matter what world one inhabits, skeptical or otherwise" (p. 7). Thus:

> **P5**: The skeptic's argument for (S_1) of AI is motivated by way of a pretheoretic commitment to internalism since, on the skeptic's analysis, warrant-conferring properties must be properties to which one has introspective or cognitive access.

Given P5, we can now ask ourselves the following question: *Why* exactly does the skeptic believe that (S_1) of AI is true? Bluntly stated, the answer seems to be something like this:

> **P6**: On the skeptic's analysis of AI, what ultimately accounts for my failure to know that I am not a BIV is that I fail to have introspective or cognitive access to properties which would indicate to me that the BIV hypothesis is false.

In other words, on the skeptic's analysis of AI, the reason as to I fail to know that I am not a BIV is that the skeptic is an *internalist* with respect to warrant; She believes that in order for

me to be warranted in the belief that I am not a BIV, I must have introspective access to an "internal mark" of some sort which would enable me to discriminate the state of affairs in which I am a BIV from the state of affairs in which I am *not*. But, the skeptic points out, I *do not* have introspective access to any such internal mark. Consequently, I do not *know* that I am not a BIV, regardless of whether or not I am actually a victim of such a scenario.

Putting Things Together

As I have already indicated, DeRose claims that it is a strength of his contextualist analysis that it successfully explains why the skeptic's AI is so persuasive in the first place. Not only that, however; As P3 – the self-containment thesis – states, DeRose believes that he is able to explain the underlying dynamics of the skeptic's AI by utilizing *solely* the epistemological and linguistic resources contained within his contextualist analysis. DeRose believes, in other words, that his contextualist analysis offers a genuinely *self-contained explanation* of skepticism, an explanation that does not appeal to any *non*-contextualist or *extra*-contextualist resources. But, crucially, is the self-containment thesis even true? Recall that the self-containment thesis is crucial to the overall success of DeRose's contextualist analysis, for if this thesis is *not* true, then DeRose's analysis can scarcely be called a *solution* to the skeptical puzzle generated by AI.

Recall DeRose's analysis of (S_1) of AI. On DeRose's analysis, what accounts for my failure to know that I am not a BIV is that, by making salient the BIV hypothesis, the skeptic raises the standards for knowledge to such a level that my belief that I am not a BIV fails to be modally stable. As P1 makes explicit, it is the fact that the skeptic places me in a high-standards context in which my belief that I am not a BIV is manifestly modally *un*stable which accounts for the truth of (S_1). But, given P1–P6, this analysis can not possibly be correct; If the results in the previous sections are correct, then on the *skeptic's* analysis, what accounts for my failure to know that I am not a BIV is *not* that, in presenting the BIV hypothesis, I am somehow forced by her into a high-standards context in which my belief that I am not a BIV fails to be modally stable, as DeRose's P1 states, but rather – as P6 and the skeptic's basic argument for (S_1) both make clear – I fail to have introspective or cognitive access to properties which would enable me to discriminate the state of affairs in which I am a BIV from the state of affairs in which I am *not*. Again; On the skeptic's analysis of (S_1), what ultimately accounts for my failure to know that I am not a BIV is *not* that, in presenting the BIV hypothesis, she somehow manipulates the semantic standards for knowledge to such a level that my belief that I am not a BIV fails to track the truth across the contextually-determined sphere of worlds (as DeRose claims), but rather that there is nothing introspectively available to me – an internal mark of some sort, something I can spot merely by focusing my attention – which would enable me to discern that I am *not* an envatted brain. For better or worse, the skeptic is, as P5 makes explicit, an *internalist* with respect to warrant; She believes that my belief that I am not a BIV is warranted only insofar as I have introspective access to properties which would serve to evidentially ground and confer warrant on my belief that I am *not* a BIV. But since, by P6, I *do not* have access to any such properties, I consequently do not *know* that I am not a BIV. Thus, it is a pretheoretic commitment to the doctrine of internalism that motivates the skeptic to affirm the boldly skeptical conclusion that I am *never* able to know that I am not a BIV regardless

of whether I am actually a victim of this scenario.[9] To be sure, as Bredo points out (2001), the skeptic's pretheoretic commitment to internalism is not at all obvious when one considers AI in the "stark, uncomplicated form" (DeRose, 1995, p. 2) in which DeRose and other contextualists present it, but it certainly stands out in the skeptic's basic argument for (S_1) which, by P5, proceeds by way of *straightforwardly internalist intuitions*.

Put slightly differently, as P4 indicates, DeRose's contextualist analysis of AI is embedded within a broadly externalist framework; On this analysis, my belief that I am not a BIV fails to count as an instance of knowledge, *not* because I fail to possess supporting evidence for this belief, *not* because I fail to have introspective access to the state, condition or property that confers warrant on this belief, but *solely* because this belief fails to satisfy the skeptically induced standard for truth-tracking that governs high-standards contexts – *solely* because, in other words, my belief that I am not a BIV fails to track the truth out to the nearest world in which I *am* a BIV. It is, thus, DeRose's putative commitment to *externalism* that is doing much, if not all, of the important epistemic work in explaining why (S_1) is true. But the skeptic will find this externalist analysis of (S_1) wholly mistaken. Blithely to say that my belief that I am not a BIV fails to count as an instance of knowledge because this belief fails to track the truth across the relevant sphere of worlds is to demonstrate a manifest failure to understand the skeptic's complaint. Either that or it is to switch the focus to an epistemic property – viz., *truth-tracking* or *modal stability*, perhaps – that is not, from the skeptic's perspective, philosophically *relevant*. To repeat; What ultimately accounts for my failure to know that I am not a BIV (on the skeptic's analysis) is that I fail to have introspective access to those properties which would enable me to discern that I am not an envatted brain. Again, by P5, the skeptic is, for better or for worse, an internalist with respect to warrant.

Furthermore, and perhaps most significantly, it is important to note that, on the skeptic's analysis of (S_1), I fail to know that I am not a BIV on the basis of what is introspectively accessible to me *not* merely in those contexts in which some set of unusually strict standards are in place, as DeRose's P1 asserts, but rather I fail to know that I am not a BIV in *any conversational context whatsoever*. As Feldman (2001), Bach (2004), and Black (2002) have recently pointed out, the leading idea behind the historically important and philosophically interesting versions of skepticism is not that we fail to satisfy some extraordinarily high standards for knowledge, but rather that, contrary to common belief, we do not satisfy even what by *ordinary* standards counts as knowledge. As Bach (2004) has put it:

> The contextualist's attempt to marginalize skeptical arguments by restricting them to skepistemic contexts ignores the fact that skepticism denies that we have knowledge even by *ordinary* standards. (p. 18)

Bach's point is that, when the skeptic raises the BIV hypothesis (or any other well-chosen skeptical hypothesis) and argues that we can not rule it out on the basis of what is introspectively accessible to us, she is not threatening to raise the standards such that *at that inflated standard* we fail to know that we are not BIV's. Rather, she is using the BIV hypothesis to show that we fail to know that we are not BIV's "even by *normal* standards governing *ordinary* contexts" (p. 18). Again, as Feldman (2001) says, non-trivial versions

[9] As Kornblith (2000) puts it, since internalism serves as the "epistemic foundation for skepticism" it is precisely the acceptance of internalism by the skeptic which ultimately leads to the conclusion that "knowledge of the external world is *impossible*" (p. 29).

of skepticism do not operate by raising the standards of knowledge. Rather, such versions of skepticism operate by getting us to see that there is a question about whether the evidence introspectively available to us enables us to "satisfy the *ordinary* standards for knowledge" (p. 80).[10] We might say, then, that the skeptic is an *invariantist* with respect to knowledge in that she believes that the standards for knowledge are *always* comparatively high, *irrespective* of context. So, on skeptical invariantism (or what Kornblith (2000) has called "full-blooded skepticism"), (S_1) is, contrary to DeRose's P1 and his contextualist analysis, *true in all conversational contexts*. But now notice; If, on the skeptic's analysis, (S_1) is true in all contexts, then by *modus ponens* on (S_2) and contrary to DeRose's P2, (S_3) is true in all contexts *as well*. Thus, on the skeptic's analysis, AI is not merely sound in high-standards contexts, as DeRose wants us to believe, but sound even by *normal* standards governing *ordinary* everyday contexts.

By way of initial summary: DeRose's contextualist explanation of why (S_1) is true appears to be mistaken. (S_1) is true *not* because the skeptic forces me into a high-standards context in which my belief that I am not a BIV fails to be modally stable, as DeRose's P1 asserts, but rather, as P6 asserts, because I can not tell "from the inside", as it were, that I am not an envatted brain. In effect, DeRose has, by P4, offered an analysis of (S_1) that proceeds *entirely by way of his externalism*. But, by P5, the skeptic is an *internalist* with respect to warrant. On the skeptic's analysis, what ultimately accounts for my failure to know that I am not a BIV is that I fail to have introspective access to properties which would enable be to discern that I am not a BIV. Moreover, on the skeptic's analysis, I fail to know that I am not a BIV on the basis of what is introspectively available to me *not* merely in those contexts in which some set of unusually strict standards are in place, as DeRose's P1 asserts, but rather I fail to know that I am not a BIV in *any conversational context whatsoever*. But now, if (S_1) is true in all contexts, then by *modus ponens* on (S_2) and contrary to DeRose's P2, (S_3) is true in all contexts *as well*. Thus, on the skeptic's analysis, AI is not merely sound in high-standards, as DeRose wants us to believe, but sound even by *normal* standards governing *ordinary* everyday contexts; On the interesting and non-trivial versions of skepticism, skepticism bites into even our ordinary everyday knowledge claims.

So, what is the upshot? Recall that, as P3 states, one of the main advantages claimed by DeRose to favor his contextualist analysis is that it functions as a *self-contained explanation* of skepticism, an analysis which successfully explains the underlying dynamics of skepticism by utilizing solely the epistemological and linguistic machinery contained within his contextualist theory. But, given that, as we have just seen, DeRose *misunderstands the full-blooded skeptic* and her reasons for accepting (S_1), it must be conceded that, contrary to P3 and the self-containment thesis, DeRose's contextualist analysis simply does *not* succeed as a self-contained explanation of skepticism. In fact, I think we can go one step further and say that, given that DeRose fails to capture the dynamics of the skeptic's reasoning underlying AI – given that DeRose's contextualist analysis *fails to explain and contain the skeptical data* – DeRose's analysis is simply *irrelevant* to the main concerns of skepticism.[11] If this is right, then one very prominent version of contextualism, DeRose's version, cannot be considered a satisfactory solution to the puzzle generated by AI.

[10] Italics added.

[11] See Johnsen (2001), pp. 392–394 for a similar conclusion.

Acknowledgments My thanks to Duncan Pritchard and Hamid Vahid for helpful comments on an earlier draft of this paper. Thanks also to Kwantlen University College for a Minor Research Grant which helped fund the research for this paper.

References

Bach, K. (2004). The emperor's new "knows". In G. Preyer & G. Peter (Eds.), *Contextualism and philosophy: On epistemology, language, and truth*. Oxford University Press. Retrieved from http://userwww.sfsu.edu/~kbach/.

Black, T. (2002). A Moorean response to brain-in-a-vat scepticism. *Australasian Journal of Philosophy, 80* (2), 148–163.

Black, T. (2006). Contextualism in epistemology. *The Internet Encyclopedia of Philosophy*. Retrieved from http://www.iep.utm.edu/c/contextu.htm.

Byrne, A. (2004). How hard are the skeptical paradoxes? *Nous, 38*(2), 299–325.

Cohen, S. (1986). Knowledge and context. *Journal of Philosophy, 83*, 574–583.

DeRose, K. (1992). Contextualism and knowledge attributions. *Philosophy and Phenomenological Research, LII*(4), 385–406.

DeRose, K. (1995). Solving the skeptical problem. *Philosophical Review, 104*(1), 1–51.

DeRose, K. (2000). How can we know that we're not brains in vats? *The Southern Journal of Philosophy, XXXVIII*(Supplement), 121–148.

Feldman, R. (2001). Skeptical problems, contextualist solutions. *Philosophical Studies, 103,* 61–85.

Greco, J. (2000). Virtue, skepticism, and context. In G. Axtell (Ed.), *Knowledge, belief, and character* (pp. 55–71). Lanham, MD: Rowman and Littlefield.

Hill, C. (1996). Process reliabilism and Cartesian scepticism. *Philosophy and Phenomenological Research, LVI*(3) 567–581.

Johnsen, B. (2001). Contextualist swords, skeptical plowshares. *Philosophy and Phenomenological Research, LXII*(2) 385–406.

Kornblith, H. (2000). The contextualist evasion of epistemology. In *Philosophical Issues, 10,* 25–32. (Special issue on skepticism edited by Ernest Sosa and Enrique Villanueva.)

Lehrer, K. (2000). Sensitivity, indiscernibility, and knowledge. *Philosophical Issues, 10,* 33–37. (Special issue on skepticism edited by Ernest Sosa and Enrique Villanueva).

Lewis, D. (1996). Elusive knowledge. *Australasian Journal of Philosophy, 74*, 549–567.

Luper, S. (2003). Introductory essay. In S. Luper (Ed.), *The skeptics*. Aldershot: Ashgate. Retrieved from http://www.trinity.edu/departments/philosophy/sluper/publications.htm.

Nozick, R. (1987). *The possibility of knowledge*. S. Luper-Foy (Ed.). Rowman and Littlefield.

Pritchard, D. (2002). Radical scepticism, epistemological externalism and closure. *Theoria, 69,* 129–161. Retrieved from http://www.philosophy.stir.ac.uk/old/staff/duncan_pritchard.htm.

Rowlands, M. (2003). *Externalism: Putting Mind and World Back Together Again*. McGill-Queen's University Press.

Sosa, E. (2003). Relevant alternatives, contextualism included. *Philosophical Studies*, 35–65.

Vahid, H. (2005). *Epistemic justification and the skeptical challenge*. London: Palgrave Macmillan.