

# KNOWING WHERE WE ARE, AND WHAT IT IS LIKE

Robert Stalnaker

[This is work in progress - notes and references are incomplete or missing. The same may be true of some of the arguments]

I am going to start with Frank Jackson's Mary, and the knowledge argument, but my main concern will not be the usual one of reconciling the case of Mary with materialism. I am interested in what the case of Mary shows about the role of phenomenal experience in providing a foundation for knowledge. After a quick review of the story, the argument, and some of the ways philosophers have responded to the argument, I will develop an analogy, exploited by one familiar strategy of response to the argument, between Jackson's puzzle and puzzles raised by John Perry and others about essentially indexical or self-locating belief. Whether or not this analogy provides a satisfactory response to the knowledge argument, I think that the similarities and differences between the two puzzles are illuminating. But getting the analogy right depends on getting the right account of essentially indexical belief, so I will spend some time sketching an account of the informational content of self-locating belief, and of the kind of fact that we state when we say who and where we are. I will conclude by looking at one reason for resisting the analogy between self-locating knowledge and knowledge of phenomenal experience, a reason that depends on an attractive, but (I will argue) untenable assumption about the role of phenomenal experience in determining what we know.

## 1. The Knowledge Argument

Everybody knows about Mary. She is a brilliant scientist who has been confined since birth to a black and white room. She knows, from reading the black and white books that line the shelves of her room, all there is to know about the physics of color, and the neurophysiology of color vision, but she has never had the opportunity to see colors. Even though she knows all the relevant physical and biological science, there is still something she does not know, something that she will learn only when she first emerges from her room, and sees colored things: she doesn't know what it is like to see colors.

Jackson's story provided the context for an argument, *the knowledge argument*, which goes roughly like this: The story seems to imply that a person might know all the relevant physical facts while remaining ignorant of certain further facts - facts about the qualitative character of visual experience. So there must be facts to be known that are not physical facts. But if there are facts that go beyond the physical facts, then materialism - the thesis that all facts are physical facts - is false.

The conclusion of the argument is that a certain metaphysical thesis is false, and most of the responses have been attempts to rebut this conclusion by reconciling, in one way or another, the thought experiment with materialism. I will review some of those responses, but my main concern will be with what the story, and some variations on it, might show us about our

epistemic relation to our experience and about the relation between our experience and our knowledge more generally.

It has been suggested that the knowledge argument is a nonstarter, since it “illegitimately draws a *metaphysical* conclusion - that physicalism is false - from an *epistemic* premise - that physically omniscient Mary would not know everything.”<sup>1</sup> But there is nothing wrong, in general, with deriving metaphysical conclusions from epistemological premises, so long as the argument is valid. And there is no mystery about how epistemological premises can have metaphysical consequences, since knowledge implies truth. *Prima facie*, at least, it is reasonable to take *facts* to be the things that are known. On this assumption, one can validly reason from an epistemological premise that Mary knows all the facts of kind K, but does not know the fact that P to the metaphysical conclusion that the fact that P is not a fact of kind K. And this seems to be roughly the form of the knowledge argument.

Of course one might dispute the assumption that it is facts that are known, arguing that what one knows is something like a fact-under-a-mode-of-presentation. What has come to be called “the Fregean response” to the knowledge argument rejects the premise that there is a fact that physically omniscient Mary fails to know. What she learns only after emerging from her room is not a new fact, but an old fact known in a new way. Mary knew all about colors under one kind of mode of presentation, but not under the mode that presents colors visually. Sometimes the point is put in terms of *concepts*: Mary, when she first sees red, acquires a new concept for a feature of the world of which she is already aware, under a different concept. I will come back to the Fregean response, but for now, I will assume that distinctions of propositional content will yield distinctions between possibilities, in some sense, and that if there is *information* that Mary lacks, when still in her room, then there are possibilities that this information would exclude.

David Lewis argued that to avoid the anti-materialist conclusion of the knowledge argument, we need a more radical solution, since he granted that distinctions of informational content bring distinctions between possibilities. He argued that unless we reject the assumption that what Mary lacks is a kind of information, then we will be committed to what he calls *the hypothesis of phenomenal information*, which is the hypothesis that

besides physical information there is an irreducibly different kind of information to be had: *phenomenal information*. The two are independent. Two possible cases might be exactly alike physically, yet differ phenomenally. When we get physical information we narrow down the physical possibilities, and perhaps we narrow them down all the way to one, but we leave open a range of phenomenal possibilities. When we have an experience, on the other hand, we acquire phenomenal information; possibilities previously open are eliminated, and that is what it is to learn what the experience is like. (Lewis 1988: 270)

---

<sup>1</sup>Alex Byrne () makes this claim, citing Horgan. He dismisses the knowledge argument with this remark, but goes on to use the Mary story to raise a different problem.

This hypothesis, since it denies that the phenomenal supervenes on the physical, “is just what minimal materialism denies.” Lewis contrasts the hypothesis of phenomenal information with what he calls *the ability hypothesis*, which claims that what Mary can acquire only from experience is not information of any kind, but only certain abilities, for example the ability to recognize colors, and to imagine seeing them.<sup>2</sup>

I think there is something right about Lewis’s ability hypothesis, but it seems to me clear that the abilities in question are *cognitive* abilities (and I don’t think Lewis would disagree). Even if learning what it is like is not, in itself, the acquiring of a certain piece of information, it still may be a precondition for acquiring a distinctive kind of information, and if there is a kind of information that Mary, without these abilities, cannot have, then we will still face the problem that the knowledge argument poses.

Part of Lewis’s persuasive case against the hypothesis of phenomenal information was the observation that while the knowledge argument was ostensibly an argument against materialism, it made no assumptions about the content of a materialist theory; if the argument works against materialism, it works equally well against any theory, including a dualist one, that takes phenomenal facts to be features of the objective world. The source of the problem is that phenomenal information seems to be, in some sense, essentially subjective. The problem is to make sense of subjective information. Are there *subjective* facts that are irreducible, not only to the physical facts, but to the facts hypothesized by any objective conception of the world as it is in itself?

## 2. Essentially indexical information

There are *indexical* or *self-locating* facts, or at least indexical or self-locating information - information, not about what the world is like in itself, but about where one is in the world - a kind of information that does not seem mysterious or metaphysically problematic despite the fact that it is irreducible to information about the objective world. Many philosophers (John Perry, David Papineau, Tim Crane, among others) have noted the analogy between Jackson's puzzle about Mary and the examples that show that knowledge and belief about the subject's place in the world are not reducible to knowledge and belief about the way the world is in itself. I want to explore the analogy, not only because I think the comparison sheds light on knowledge of phenomenal experience, but also because I think it might help us to get clearer about essentially indexical knowledge and belief themselves, and about knowledge generally.

There is a familiar range of examples - stories told by Hector-Neri Castaneda, John Perry, David Kaplan, David Lewis, among others to illustrate the fact that our knowledge of ourselves and our place in the world - knowledge about what is happening *now*, or *here*, or to *me* - cannot be reduced to impersonal objective knowledge about what the world, as it is in itself, is like. There is John Perry's "shabby pedagogue" who sees himself in the mirror and forms beliefs about the

---

<sup>2</sup>Lewis, “What experience teaches.” See also Nemirov

person he sees without realizing that it is himself. There is the case of the person who knows, at noon, that the meeting starts at noon, but does not get up to go to it, since she does not know that the meeting starts *now*. There are various stories about amnesiacs, and then there is David Lewis's story of the two omniscient gods, each of whom knows exactly what possible world they are in, but not which of the two gods he himself is.

The moral of these stories is supposed to be that no amount of information about what the objective world is like will by itself tell you where you, the possessor of that information, are in that world. Just as the case of Mary suggests that there are facts that an objective, physical description cannot capture, so the cases of essentially indexical belief suggest that there is a special kind of fact about who and where we are that objective description cannot capture. Few will be tempted to respond to this problem by enriching or refining our conception of the impersonal objective world - by *objectifying* the self - but how should we understand such information?

If our question is, exactly when are self-locating statements and beliefs true, then the answer is clear and unproblematic: A statement of the form "I am F" is true, when said or thought by X, if and only if X is F. So if David Kaplan says "my pants are on fire", what he says is true if and only if David Kaplan's pants are on fire at the time at which he says it. But this answer does not tell us what the *contents* of the statements are - what *information* they convey, what belief they express, what kind of *fact*, if any, they state. One influential answer to this question, defended by David Lewis, is this:

If the contents of ordinary beliefs about objective facts can be represented by sets of possible worlds, then the contents of self-locating beliefs can be represented by sets of *centered* possible worlds - possible worlds each with a designated time and person or object. Since "I am sad" is true if and only if it is said or thought by a sad person, its content will be represented by the set of centered possible worlds that have a sad person at their designated center. The content of "the meeting is about to start" will be a set of centered worlds at which the meeting in question takes place soon after the time that is designated as the center.

There are a number of problems with this account of self-locating content. First, it identifies contents that ought to be distinguished. What I believe when I believe that I was born in New Jersey is something about myself, something different from what my fellow New Jersey natives believe about themselves. What I tell the waiter when I tell him that I will have the mushroom souffle is different from what you tell the waiter if you decide to have the same thing. But on the centered worlds account, our respective beliefs and statements have the same content. Second, and more important, the account distinguishes contents that ought to be identified. If Rudolf Lingens tell you that he is sad, or that he is Rudolf Lingens, and you understand and accept what he says, then it seems that the information you acquire is the same information that he imparted. But you do not, of course, thereby ascribe the property of being sad, or of being Rudolf Lingens to yourself. Third, it is not clear how the account deals with the wider class of context dependent

expressions - demonstratives and second person pronouns, for example. When I believe that you are sad, what property am I self-ascribing? If I say "your pants are on fire," do I say the same thing as I would say if I said, "I am talking to person whose pants are on fire"?

To set up an alternative to Lewis's account that I think gives more adequate representation of self-locating content, let me start with a slightly different question: not, "what is the content of a particular indexical statement or belief?" but "how should a person's state of belief as a whole be represented so that it includes his or her self-locating beliefs?" On the old view that ignored the phenomenon of self-location, a belief state was represented by a set of possible worlds - the doxastically accessible worlds. The Lewis view simply replaced this set of worlds with a set of centered worlds. In both cases, the representation of a belief state is abstracted from the believer. But the lesson of the phenomenon of self-locating belief is that we cannot give an adequate representation of a state of belief without connecting the world as the subject takes it to be with the subject who has the beliefs. What we want to represent is the state of belief that a particular individual *x* is in at a particular time *t* in a particular possible world *w*. When we represent the way this individual locates himself in the world as he takes it to be, we need to include the information about who it is who is locating himself there, and we need to link the world as the believer takes it to be to the world in which the believer takes the world to be that way. I will use what I will call a set of *labeled* or *indexed* possible worlds to represent a state of belief. This is like a set of centered worlds except that the designated center is indexed to a particular individual, time and world - the individual who is in that belief state at that time in that world.

By linking the world according to a believer to the world in which the believer has the beliefs, we can represent and compare the beliefs of more than one person and the labeling also facilitates the representation of iterated beliefs: with what one person believes about what another believes. This is crucial for the representation of the communication of self-locating information, which is difficult to make sense of on a model that uses sets of centered worlds to represent the information. We want, for example, to connect the belief Alice expresses when she tells Bert that she is Alice with what Bert comes to believe when he accepts what she says. When we represent iterated and shared beliefs, the account will yield possible worlds with multiple indices. If we model a *conversational context* by a set of possible worlds that represents the common ground - presumed mutual beliefs of the participants in the conversation, the account will yield a label and a link for each participant. The common ground in the context will represent not only what the participants presume the world to be like, but where they locate themselves and each other in the world.

It is true in general that what is said (self-locating or not) is said in a context, against a background of shared information that includes information about the context itself. Possible worlds compatible with the context - the shared background information - are possible worlds in which the participants in the conversation exist, and are having the conversation that they are having. This set of (indexed) possibilities is the set of live options that they intend to distinguish between with their speech acts. In making an assertion in such a context, one expresses a

proposition (which might be represented by the set of possible situations in which it is true), which, if it is accepted, changes the context by eliminating the possibilities in which it is false. In some cases, the labels will be irrelevant to the information that is conveyed, that is, irrelevant to the way the possibilities are distinguished by the content of the speech act. Or, it may be that the indices are relevant only as a means of determining the information the speaker intends to convey, and not to the information itself (Suppose self-identification is not at issue - you know perfectly well who I am - but I use the first person pronoun to tell you that I was born in New Jersey. You take away from the conversation the objective information that Bob Stalnaker was born in New Jersey.) In such a case, we can detach the information from the context in which it was expressed, or from the situation in which it is believed. That is, we can identify the content of what is said or thought (the way it distinguishes between possible worlds) independently of the fact that it was something that was said or thought on that particular occasion. But sometimes, the information - the way a speech act or thought distinguishes between the possibilities that define the context - essentially involves the labels (it is information about the participants in the conversation as participants in the conversation), and so cannot be detached from the context in which it is expressed or thought. In this kind of case, I will say that the information is *contextual* or *subjective* information. So suppose you didn't know who I was, and what I told you was not that I was born in New Jersey, but that I was Bob Stalnaker. You would presumably learn *something* about the objective world from what I told you, together with what you already knew and what you observe, but there is not a piece of information that is the content of what I told you that you can simply add to your stock of beliefs about the objective world.

The point is not simply that a proposition can be about a context in which something is said, or about a person's location in the world. A piece of information about a speech act, or about a judgment or a belief can be an ordinary objective context-independent proposition about the world as it is in itself. It might be, for example, an objective fact about the centerless world that David Kaplan said, or believed, on January 14, 1975 at 2:15 PM (PST), that his pants were on fire. The point about essentially contextual information is that sometimes the content of what is expressed or believed in a context is not detachable from the context in which it is expressed or believed. Suppose Alice points to a bomb and says "that is about to explode!" What she said is true if and only if she is pointing (on the particular occasion in question) at something that is about to explode. Another statement might have been made *about* Alice's demonstration and statement - say the day before that the next day, Alice would point at a bomb in such and such circumstances, and that it would soon after explode. This statement made the day before might be true in the same possible worlds as Alice's statement, but it doesn't seem right to say that it conveys the same information, or has the same content.

### **3. Developing the analogy**

Is the knowledge that Mary lacks while still in her room knowledge that is essentially contextual in this way? It is knowledge that, after she gets it, is naturally expressed with a demonstrative ("Now I know that seeing red is like *this*"), and her relation to what she is demonstrating (the type of experience she is having or recalling) seems to be essential to the character of the information it

is used to express in the way that it is in the case of essentially indexical information. But the analogy, if it implies that Mary is not ignorant of any relevant fact about the world as it is in itself, may seem strained. Before considering one reason why it might seem strained, let me sketch a variation on the story about Mary that I hope will make clearer exactly what the analogy is supposed to be.<sup>3</sup> (This is a variation I discussed in an earlier paper on Thomas Nagel's objective self)

Suppose that Mary, still in her room, is told that she will be subjected to the following experiment. She will be shown either a red or a green star, to be chosen by the flip of a coin, and she is told in great detail the exact circumstances of the two possible scenarios. So given her extensive knowledge of neurophysiology and color science, she knows that when the experiment is performed, she will be in the presence of a star with one of two specific light reflectance properties, and will be in one of two specific brain states. Both before and after the experiment is performed, there are two possible worlds compatible with Mary's knowledge - call them worlds R and G. As it happens, the red star is chosen, so she is in fact in possible world R.

In a sense, after the experiment is performed, Mary knows what it is like to see red, although not under that description. "Now I know," Mary says, "either what it is like to see red, or what it is like to see green. I just don't know which it is, since for all I know, *this* experience could be the experience of seeing red, or of seeing green."

What changed about Mary's epistemic situation, when she was shown the star, is that she was then a position to represent subjective information about *this* experience, just as Alice, at the scene of the impending explosion was in a position to represent, in her speech and her thought, the contextual information that *that* bomb is about to explode (that is, will explode soon after *now*). Note that Mary's situation is *not* being compared with the situation of amnesiacs or people who do not know what time it is, or where they are. Rather, the analogy is this: Mary is like the person who is not in a position to know a certain piece of contextual information because she is not in the relevant context. Mary's situation with respect to color (when she is still in her room) is being compared with her situation with respect to Alice, whose warning about the bomb took place far away, and at a different time. Alice knew, as she put it at the time, that *that* is a bomb and that if we don't defuse it, it will go off five minutes from *now*. Mary might have known that there was a bomb at a certain place in the Rose Garden at 6:47 PM, being demonstrated by Alice, and that if it is not defused, it will explode at 6:52 PM. But this is not what Alice knew, and Mary is not in a position to know what Alice knew. She would have had to be there. Had she been there, she would then have been in a position to represent the thought that *this* is the bomb, and that it will go off five minutes from *now*.<sup>4</sup>

---

<sup>3</sup>I discuss this variation in an earlier paper, "Thomas Nagel's Objective Self"

<sup>4</sup>Note to Chalmers's response to Perry in PPR

#### 4. A problem with the analogy

If you buy this analogy, then you can explain the knowledge Mary lacks, when still in her room, and you can understand that knowledge in terms of the elimination of possibilities, without invoking the hypothesis of phenomenal information that explains phenomenal information in terms of distinctions between *objective* possibilities. But, the fan of qualia will protest, the analogy won't fly. There is nothing essentially indexical or demonstrative about the information Mary lacks. Let's assume she had, while still in her room, a name, "ph-red" for red-type qualitative experience, though she didn't know what it was like to have a ph-red experience.<sup>5</sup> Then, when she is shown the red star, she wonders whether *this* experience is ph-red. But Mary could then coin another name for her experience. Suppose, following John Perry, she names it "wow". Now she can express her question whether *this* is ph-red as a question about a context-independent, objective proposition: is it true that wow = ph-red?

But note that the same kind of maneuver could be made in the Rose Garden bomb scenario. Mary, since she doesn't know what time it is when she arrives, says, "I hereby dub the time five minutes from now, "pow". "I know the bomb will explode at 6:52," she says, but will it explode at pow?"

But (the objector continues) the cases seem different in this way: when Mary named the time "pow" she didn't know what time she was naming, and so didn't know what objective proposition she was asking about. But when she saw the red star, and named her experience "wow", she knew what she was naming, since she was *acquainted* with the experience. But I want to question the assumption that there is something - phenomenal experience - that has both an autonomous place in an external conception of the world as it is in itself and also this kind of distinctive epistemic role. I will conclude today by looking at an intuitively attractive and widely shared assumption about the relation between knowledge and experience that I think is one of the sources of resistance to the kind of austere account of subjective information that I am promoting.

The assumption, which I will call the principle of phenomenal indistinguishability, is this:

If two possibilities are epistemic alternatives for a knower at a time (that is, both are compatible with his or her knowledge) then they are *phenomenally indistinguishable* to the knower at that time.

The empiricist holds that the evidential basis for all our knowledge is experience - phenomenal experience. Whatever the right story to tell is about how our knowledge can get beyond our experience, it seems natural to assume that we at least can know that the quality of our immediate experience is the way that it in fact is. But our story about Mary in the coin flip scenario seems to conflict with this natural assumption.

---

<sup>5</sup>"ph-red" is Bill Lycan's name for phenomenal red.

Mary has just seen a red star, but is ignorant of whether it is red or green. The two possible situations we used to represent her ignorance were physically different: World R is the actual world, while in nonactual world G she is presented with a green star, with all the optical and physiological consequences that that would have had in the actual world. (Since Mary knows all the general optical and physiological facts, these facts will hold in all worlds that are epistemically possible for her.) But it seems intuitively that these two possible worlds are phenomenally as well as physically different. If the flip of the coin had been different, and Mary had been shown the green star instead, things would have looked very different to her. Nevertheless, she still doesn't know which of the two situations she is in.

If it is assumed that epistemic alternatives must be phenomenally indistinguishable, then world G cannot be compatible with Mary's knowledge. To account for Mary's ignorance of which color she has seen in this case, one will be required to suppose that there is a different possible world, G\* which is *physically* just like the world G, but *phenomenally* just like the world R, and if we accept that, we have bought the hypothesis of phenomenal information.

While I think we should reject the principle of phenomenal indistinguishability, it is a seductive assumption: Lewis himself gives an account of knowledge that commits him to it.

What we know, on Lewis's account of knowledge, is what is true in all of the possibilities that are uneliminated by our evidence, and evidence is identified with experience. Lewis gave a characteristically sharp and clear formulation of what it means for a possibility to be eliminated by evidence: "I say that the uneliminated possibilities are those in which the subject's entire perceptual experience and memory are just as they are. . . . A possibility *W* is *uneliminated* iff the subject's perceptual experience and memory in *W* exactly match his perceptual experience and memory in actuality."<sup>6</sup> Lewis, in the context in which he gave this definition, is mainly concerned with the problem of how, if knowledge is understood in terms of the elimination of possibilities by experience, we are able to answer the skeptic. He defended a contextualist account of knowledge that permits us, in context, to ignore certain possibilities that are *not* eliminated by experience. But on Lewis's account, we will at least know, *in any context*, that the possibilities excluded by our experience - possible situations in which our experience does not match our actual experience - are possibilities that are incompatible with our knowledge. But all of these possibilities will be possibilities in which our experience has whatever essential properties our actual experience has. That is, Lewis's account of knowledge implies that even in the most skeptical context, we will know the essential nature of our experiences - "know exactly what they are . . . in an uncommonly demanding and literal sense of 'knowing what'."<sup>7</sup>

---

<sup>6</sup>Lewis, "Elusive Knowledge"

<sup>7</sup>Lewis, "Should a materialist believe in qualia? In this paper, Lewis argues that we should reject what he calls "the identification thesis," which is the thesis that we know the essential nature of our qualia, simply by experiencing them. But I am arguing that Lewis's account of knowledge commits him to this principle.

Suppose we reject the principle of phenomenal indistinguishability and grant that there are possibilities that Mary, in the coin flip scenario, cannot exclude in which her phenomenal experience is different from what it actually is. The two epistemic alternatives that Mary cannot distinguish between, after having seen the red star, are the same two that she could not distinguish before seeing it. So what *were* the epistemic consequences, for Mary, of leaving her room, and seeing color for the first time - what did she learn? It is tempting to say that what she did was to acquire a new *concept* for the qualitative property that she came to exemplify. She already had one concept of that property, before leaving her room, and after she was presented with the red star, she had two concepts, without knowing that they were concepts of the same property. The second concept Mary acquired was a *pure phenomenal concept*.<sup>8</sup> It is tempting to say this, and there must be something right about it, but I don't think these creatures of darkness - concepts - will help us to clarify what is going on. We can't avoid our problem by introducing an additional layer between the knower and the features of the world (or in this case, features of the subject) that are known. (The main attraction of concepts, I think, is that they facilitate equivocation between the vehicles of representation - the linguistic or mental objects or features that do the representing - and the meaning or content of the representation.)

So what are concepts? Some think of them as something like mental words - words of a language of thought. Presumably, their meanings are essential to them - to grasp a concept is to know its meaning. Or perhaps the concept *is* the meaning. But what is the meaning of the pure phenomenal concept that Mary acquired - the one that (following John Perry) she used the word "wow" to express? More specifically, is the concept that Mary acquired in the actual world (the world in which she was shown a red star) the same concept as the one she acquired in the alternative world in which she was presented with a green star? We know that Mary would have had a different kind of experience if she had been shown the green star, but would the *concept* that she acquired have been different? There are problems with each answer to this question.

If the two "wow" concepts are the same, then Mary knows what concept she is expressing, but that concept is only contingently connected to the experience itself, so it is not really a pure phenomenal concept, but a descriptive or perhaps demonstrative concept, identified by Mary's contextual situation, and her knowledge there, rather than by the intrinsic character of the experience. On the other hand, suppose the "wow" concepts expressed in the two possible worlds are different from each other. Suppose they are object-dependent concepts, identified by the phenomenal properties themselves - the qualia - to which they in fact refer. Then they might reasonably be called pure phenomenal concepts, but we will have to conclude that Mary does not know which of the two concepts it is that she is thinking. Her thought, on this account, will violate what Paul Boghossian calls a principle of *epistemic transparency*. On this account, we have a direct connection between the content of Mary's thought and the qualitative character of the experience that the thought is about (a connection that perhaps constitutes her being *acquainted* with the character of the experience), but the cost is that we must say Mary is only

---

<sup>8</sup>note to Chalmers.

indirectly connected to the content of her thought.

The argument that, on this hypothesis about Mary's concept, she does not know the content of her thought is this: Mary does not know whether she is in world R or G; the concept she is expressing is different in world R than it is in world G; therefore, she does not know which of the two concepts she is expressing. The premise that she does not know whether she is in world R or G is based, not on a theoretical argument, but on a pretheoretic judgment about the case. It is not that a certain conception of knowledge implies that she does not have this knowledge - it is that the case is one in which it seems intuitively clear that Mary does not know whether the star is red or green. The dilemma would be resolved if we accepted the principle of phenomenal indistinguishability, and the hypothesis of phenomenal information that comes with it, but the introduction of conceptual intermediaries will not help us to avoid this conclusion.

If we reject this principle, do we have to say that Mary does not really know what qualitative character her experience has, even while she is having it? The notion of knowing what or who something or someone is is notoriously problematic. Some have argued that we never really know, about things in the external world, what they are in themselves - never know what they are in "an uncommonly literal and demanding sense", but phenomenal experience was supposed to be different. It is presented directly; it is supposed to be its own mode of presentation.<sup>9</sup> Perhaps our knowledge of the internal world is as indirect as our knowledge of what lies beyond it. Or more plausibly, perhaps we need a better conception of what it is for knowledge to be direct.

---

<sup>9</sup>Note to Levine, Loar