

# Conceptual mastery and the knowledge argument

Gabriel Rabin

Published online: 22 February 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

**Abstract** According to Frank Jackson’s famous knowledge argument, Mary, a brilliant neuroscientist raised in a black and white room and bestowed with complete physical knowledge, cannot know certain truths about phenomenal experience. This claim about knowledge, in turn, implies that physicalism is false. I argue that the knowledge argument founders on a dilemma. Either (i) Mary cannot know the relevant experiential truths because of trivial obstacles that have no bearing on the truth of physicalism or (ii) once the obstacles have been removed, Mary can know the relevant truths. If we give Mary the epistemological capabilities necessary to draw metaphysical conclusions about physicalism, she will, while trapped in the black and white room, be able to know every truth about phenomenal experience.

**Keywords** Concepts · Concept possession · Conceptual mastery · Knowledge argument · Phenomenal concept · Physicalism · Phenomenal concept strategy · Dualism · Modal rationalism · Consciousness · Phenomenal experience

## 1 Introduction

According to Frank Jackson’s famous knowledge argument (Jackson 1982, 1986), Mary, a brilliant neuroscientist raised in a black and white room and bestowed with complete physical knowledge, cannot know certain truths about phenomenal experience. For example, she can’t know what it’s like to see red. This claim about knowledge, in turn, implies that physicalism is false. I argue that the knowledge argument founders on a dilemma. Either (i) Mary cannot know the relevant experiential truths because of trivial obstacles that have no bearing on the truth of

---

G. Rabin (✉)  
Department of Philosophy, UCLA, 405 Hilgard Ave,  
Los Angeles, CA 90095-1451, USA  
e-mail: grabin@humnet.ucla.edu

physicalism or (ii) once the obstacles have been removed, Mary can know the relevant truths. If we give Mary the epistemological capabilities necessary to draw metaphysical conclusions about physicalism, she will, while trapped in the black and white room, be able to know every truth about phenomenal experience.<sup>1</sup>

## 2 Physicalism

Physicalism requires, at the least, that the physical state of our world necessitates the complete state of our world. The knowledge argument purports to refute this supervenience claim. We operationalize physicalism as follows:

Physicalism =<sub>def</sub> Every austere physical duplicate of the actual world is a duplicate simpliciter.<sup>2</sup>

An *austere physical duplicate* of the actual world is a metaphysically possible world at which every austere physical proposition true at the actual world is true. A duplicate simpliciter of the actual world is a world at which every proposition true at the actual world is true. An *austere physical proposition* (or *physical proposition*) is a proposition composed of concepts taken only from physics, mathematics, and logic, or which is expressible using only vocabulary taken from physics, mathematics, and logic. A *truth* or *fact* is a true proposition. I assume a fine-grained, roughly Fregean treatment of propositions.<sup>3</sup> On this approach, the proposition expressed by ‘Hesperus is bright’ is not the same proposition expressed by ‘Phosphorous is bright’. Concepts compose propositions in a manner similar to that by which words compose sentences.

## 3 The knowledge argument

Here’s our quick and dirty starting version of the knowledge argument.

- (1) Knowledge of all the austere physical information does not put one in a position to know all the information.
  - (2) If (1), then physicalism is false.
- 
- (3) Therefore, physicalism is false.

<sup>1</sup> Phenomenal experiences are associated with raw experiences or sensations. Some examples include the sensation associated with a pain, the color-experience associated with seeing the sky (a “blue experience”), or the sensation one has when the back of one’s neck itches. In the words of Nagel (1974), there is something “it is like” for the subject of a phenomenal experience. I use the terms ‘phenomenal’ and ‘experiential’ interchangeably.

<sup>2</sup> This definition is equivalent both to (i) the truth of every proposition supervenes on the truth of the austere physical propositions and to (ii) the true austere physical propositions necessitate every true proposition.

<sup>3</sup> This assumption is dialectically fair—it *helps* the knowledge argument, which I plan to argue against. For example, this assumption makes it much more difficult to launch the “same proposition, different mode of presentation” defense against the knowledge argument (cf. Byrne (2002), Tye (1995)). I’ll argue that Mary can know the relevant propositions, even on a fine-grained approach.

The argument is clearly valid. But is it sound? Neither premise is obvious; both have been coherently denied by physicalists.<sup>4</sup> Support for premise (2) comes from theorizing about modality. Support for premise (1) comes from Jackson's (1982, 1986) famous thought experiment involving Mary.

Jackson asks us to envision Mary, a "brilliant neuroscientist who is... forced to investigate the world from a black and white room via a black and white television monitor. She... acquires... all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red', 'blue', and so on." Mary has complete physical information, i.e. she knows every austere physical proposition, and every proposition that such knowledge places one in a position to know. Jackson asks whether Mary learns anything new when she emerges from the black and white room and sees a ripe red tomato for the first time. He concludes that "it just seems obvious that she will learn something about the world and our visual experience of it." Therefore, knowledge of all the austere physical information does not put one in a position to know all the information—i.e. premise (1) is true.<sup>5</sup>

Premise (2) gets support from the desire to avoid mysterious unexplained brute necessities and from *modal rationalism*, according to which there are important (potentially constitutive) connections between modality and epistemological notions like "in a position to know", "a priori", and "conceptual truth".<sup>6</sup> The general physicalist idea that "all information is physical information", knowable on the basis of complete austere physical information, also supports premise (2).

#### 4 The missing concept reply

The knowledge argument relies on a specific type of inference—an inference from an epistemic gap, the lack of knowability of a *target set* of propositions on the basis of a *base set*, to a metaphysical gap, the lack of necessitation from the base set to the target proposition(s). In the knowledge argument, the base set is the set  $\mathcal{P}$  of all true austere physical propositions. The target set contains, at the minimum, one proposition about phenomenal experience.

The knowledge argument, as stated, does not stand a chance. It is a mistake to conclude that a base set of propositions does not necessitate a target proposition from the fact that someone who knows the base propositions is not in a position to know the target. An agent might know the base propositions but not be in a position to even *possess* the concepts in the target proposition. As a result, the agent won't be in a position to know the target propositions—because he or she can't even consider them! But an epistemic gap of this nature shows nothing about necessitation. The following example demonstrates why. Imagine that Mary's sister Jane knows all the

<sup>4</sup> Premise (1) has been denied by so-called "type-A" physicalists, premise (2) by so-called "type-B" physicalists. This terminology originates in Chalmers (2003).

<sup>5</sup> Enthymematic here is the claim that if Mary can't know, no one can.

<sup>6</sup> Jackson (1998) is the most sustained argument to date for modal rationalism. Cf. also Chalmers (forthcoming), Chalmers and Jackson (2001), Lewis (2002), Levine and Trogon (2009), and Schwarz (2007).

information in  $\mathcal{P}$ , all the information about gender, and all the information about who is married to whom. But Jane does not possess, and is not in a position to possess, the concept BACHELOR.<sup>7</sup> Jane knows that Ursula is a woman. But she does not know, and is not in a position to know, that Ursula is not a bachelor. If we infer from this epistemological gap to a metaphysical gap, we'll conclude that the fact that Ursula is a woman does not necessitate that she is not a bachelor. But clearly this is false. Ursula's being a woman *does* necessitate that she is not a bachelor.

The lesson is that if we're going to infer from an epistemic gap between two propositions (or sets of propositions) to a lack of necessitation, and we test for the epistemic gap by checking what hypothetical agents could come to know, we must require that these agents possess the concepts in the target proposition, or at least are in a position to possess them.<sup>8,9</sup> The fact that Jane, who does not possess the concept BACHELOR, can't know that Ursula is not a bachelor despite knowing that Ursula is a woman, shows nothing about a lack of necessitation between the two propositions.

In the case of Mary, it is plausible that Mary does not possess the relevant experiential concepts. For example, consider the concept RED<sub>exp</sub>. RED<sub>exp</sub> is the phenomenal, experiential, concept of red. It is the concept tied to looking like *this* (where a red patch is demonstrated), not the concept tied to light of wave-length 700 nanometers, or to surfaces that reflect light of that wave-length. Tomatoes, fire trucks, and strawberries are red, but they cause red<sub>exp</sub> sensations.<sup>10</sup> The fact that Mary has never had a red<sub>exp</sub> experience, cannot imagine a red<sub>exp</sub> sensation, and could not recognize a red<sub>exp</sub> sensation as red<sub>exp</sub> were she to have one, supports the claim that Mary does not possess the concept. With the "concept possession requirement" in place, the physicalist has an easy reply to the knowledge argument—deny premise (2). It is foolish to conclude, from the fact that Mary's knowledge of base physical propositions does not put her in a position to know target experiential propositions, that the base does not necessitate the target. Because she does not possess the concepts that compose the target propositions, Mary cannot consider them, let alone know them. Only an inability to know, for an epistemic agent that possesses (or is in a position to possess) all the concepts in the target propositions, has any hope of yielding the desired metaphysical consequences.

## 5 Moving forward

Faced with the missing concept reply, the advocate of the knowledge argument must either abandon the argument or modify the epistemological relationship used in it.<sup>11</sup>

<sup>7</sup> I use small capitals to denote concepts, e.g. the concept BACHELOR.

<sup>8</sup> In fact, I think that much more is required in terms of concept possession. Cf. Rabin (manuscript).

<sup>9</sup> Byrne (2006), Hellie (2004, p. 348), Stoljar (2005, pp. 474–475), and Tye (2000, pp. 17–18) all make similar points. They all note that one cannot draw any substantive conclusion from the Mary case if Mary does not possess experiential concepts and cannot even consider the experiential propositions in question.

<sup>10</sup> All tomatoes mentioned in this paper are ripe tomatoes with brilliant red skins.

<sup>11</sup> The missing concept reply has appeared in the literature before (e.g. Hellie (2004), Loar (1990), Stoljar (2005), and Tye (2000)).

This epistemological relationship has two desiderata. First, its absence must be good evidence for a lack of necessitation.<sup>12</sup> Second, this epistemological relationship must fail to hold between the austere physical propositions  $\mathcal{P}$  and certain experiential propositions. Evidence for this failure will come from considering what Mary can and cannot know.

I introduce a technical term, *implication* (verb form: *implicate*), to express an epistemic relationship that might meet the desiderata. ‘Implicate’ allows us to express the following schema for the knowledge argument:

- (1)  $\mathcal{P}$  does not implicate some proposition E about phenomenal experience.
  - (2) If  $\mathcal{P}$  does not implicate E, then  $\mathcal{P}$  does not necessitate E, and physicalism is false.
- 
- (3) Therefore, physicalism is false.

Substituting for ‘implicate’ here yields different versions of the knowledge argument. A candidate notion of implication meets the two desiderata above if and only if it validates both premises. One upshot of this paper is that finding such a conception of epistemological implication is difficult. All the candidates I consider founder on one premise or the other. For example, in the previous section I argued that any conception of implication according to which ‘B implicates T’ is equivalent to (or even entails) “knowledge of B puts one in a position to know T” fails to validate the second premise. One might not know because one does not possess the concepts in T, but this shows nothing about either necessitation or physicalism. However, that discussion yields an obvious suggestion for what implication might be. Perhaps ‘B implicates T’ means that knowledge of B, given that one possesses all concepts in T, puts one in a position to know T. In the next section, I argue that this suggestion fails.

---

<sup>12</sup> A question that may have occurred to the reader is whether we are *ever* justified in inferring from the lack of knowability of a target proposition from a base set to the fact that the base does not necessitate the target. This issue becomes more pressing when one recognizes the obvious counterexamples. Even if one possesses all the relevant concepts, knowing that there is water in the glass does not put one in a position to know that there is H<sub>2</sub>O in the glass. Yet the fact that there is water in the glass does necessitate that there is H<sub>2</sub>O in the glass. The worry can be addressed and the obvious counterexamples avoided. The modal rationalist advocates of the general strategy of inferring from an epistemic to a metaphysical gap (e.g. Jackson (1998), Chalmers (1996), Chalmers and Jackson (2001)) do not endorse the general claim that, for any propositions B and T, if knowing B does not put one in a position to know T, B does not necessitate T. The inference to a metaphysical gap is valid only for a special class of base propositions. There is insufficient space for a full treatment here. One strategy is to require that, in order to draw conclusions about necessitation from facts about epistemic implication, the base propositions must be *semantically neutral*—i.e. not “twin-earthable”. Famously, Putnam (1975) showed that ‘water’ is not semantically neutral. It is plausible that the vocabulary of fundamental physics, the vocabulary in which the austere physical description  $\mathcal{P}$  is written, is semantically neutral. Anything that acts like an electron *is* an electron, whereas something can act like water and be XYZ, not water. Even if the fundamental vocabulary of physics is not semantically neutral, this merely introduces an epicycle on the dialectic (cf. Chalmers (1996, pp. 135–136)). Another important difference between the proposition that there is water in the glass and the proposition  $\mathcal{P}$  under consideration in the knowledge argument is that, according to the physicalist,  $\mathcal{P}$  necessitates all the facts—it leaves the truth value of no proposition unsettled. The proposition that there is water in the glass has no such pretensions.

## 6 The conceptual mastery reply

Some philosophers have pointed out that Mary, raised in her black and white room, *can* possess the relevant experiential concepts.<sup>13</sup> She can acquire them from her co-workers, who have seen ripe red tomatoes and the big blue sky. Mary won't have as good a grasp of these concepts as her co-workers. After all, Mary has never experienced red<sub>exp</sub>. But Mary will at least possess the relevant concepts and be able to grasp propositions that contain them.<sup>14</sup> However, it seems that Mary's acquisition of these concepts from her co-workers won't help her learn the missing experiential information. If so, then Mary can possess all concepts in the target experiential propositions, know all the austere physical propositions, and yet still not be in a position to know the experiential propositions. The knowledge argument can overcome the objection raised in the previous section and adapt to incorporate its lesson, which is that the relevant epistemic agents must possess the concepts in the target proposition.

In this section, I argue that we need additional constraints on implication. We should not infer from "knowledge that B does not put someone who possesses all the concepts in T in a position to know that T" to the fact that B does not necessitate T. Possessing the concepts in T is not enough; we must require more of the epistemic agent.

There are (at least) two different levels of understanding a concept. The first is *concept possession*. If one is able to grasp propositions that contain the concept, or think contents a component of which is the concept, then one possesses the concept. Concept possession is easy to obtain. Burge's (1979) Alfred possesses the concept ARTHRITIS, but Alfred does not know what every doctor knows—that arthritis affects joints, not limbs. Some of the doctors have a more sophisticated and demanding relation to the concept ARTHRITIS—they have *conceptual mastery*. The proverbial "experts", to which laymen defer, often possess conceptual mastery. One can possess a concept and still be grossly wrong about its extension, and even about constitutive a priori truths regarding the concept. Conceptual mastery is less tolerant of such errors.<sup>15</sup>

<sup>13</sup> Ball (2009) and Tye (2009), among others, make this point.

<sup>14</sup> A full defense of the claim that Mary can possess the relevant concepts in this way would take us too far afield. It is certainly the consensus view in philosophical research on content. Here's a brief argument for the claim. If Mary thinks that the sky causes red<sub>exp</sub> sensations, her thought is *wrong*—the sky causes blue<sub>exp</sub> sensations, not red<sub>exp</sub> ones. This is so because her thought has the same truth-conditions the thought we would express with 'the sky causes red<sub>exp</sub> sensations' does. It has the same truth conditions because it contains the same concepts, which contribute to the truth-conditions of the complete thought. Mary's concept has the same veridicality conditions our concept RED<sub>exp</sub> does. Therefore, Mary thinks with the same RED<sub>exp</sub> concept the rest of us do. Here's another brief argument: If Mary does not think with our RED<sub>exp</sub> concept, she thinks with some other idiosyncratic concept (perhaps her RED<sub>exp</sub> concept is similar to our BLUE<sub>exp</sub> concept). Interpreting her charitably, her idiosyncratic thought is most likely correct. But her thought is not correct; she is wrong. Therefore she does not think with another concept.

<sup>15</sup> The distinction between concept possession and conceptual mastery also appears in Greenberg (manuscript, 2009) and Bealer (2002). Of course, the distinction has its roots in the seminal discussion of Burge (1979). Mark Greenberg deserves credit for stressing to me the importance of the distinction and the pitfalls into which philosophers can fall when they ignore or pay insufficient attention to it. In several places (Ibid.), Greenberg uses an argumentative strategy similar to the one I adopt here. His use pre-dates

Once we have distinguished these two levels of understanding a concept, we can improve our grip on the connection between epistemic gaps and absence of necessitation. We said above that when inferring from lack of implication to lack of necessitation, only lack of knowability for agents who possesses all concepts in the target proposition is relevant. We can do one better: only lack of knowability for agents who have mastery of all concepts in the target proposition is relevant. The following example demonstrates why.

Imagine again Mary's sister Jane. By engaging with her co-workers, Jane comes to possess the concept *BACHELOR*. She knows that Ulysses, who is married, is not a bachelor, and that Achilles is a bachelor. But Jane is not a master of the concept *BACHELOR*; she thinks unmarried women can be bachelors. Jane has complete knowledge of  $\mathcal{P}$  and knowledge of the gender and marital status of every person. Jane knows that Ursula is a woman, and that Ursula is not married. Unlike the previous case, because Jane now possesses the concept *BACHELOR*, she can consider the proposition that Ursula is not a bachelor. But because of her erroneous view that women can be bachelors, Jane will not come to know that Ursula is not a bachelor. If we don't require, in the inference from an epistemic to a metaphysical gap, that the relevant epistemic agents have conceptual mastery (i.e. if mere concept possession is enough), then we will be forced to conclude, on the basis of Jane's inability to know that Ursula is not a bachelor, that Ursula's status as a female does not necessitate that she is not a bachelor. But that conclusion is preposterous. Thus, we must require, if we are to infer from a lack of implication to a corresponding lack of necessitation, that the epistemic agents have mastery of all concepts in the target proposition(s).<sup>16</sup>

The conceptual mastery requirement offers a reply to the knowledge argument.<sup>17</sup> Mary, despite coming to possess the concept *RED<sub>exp</sub>* via interaction with her co-workers, does not have conceptual mastery. In order to have conceptual mastery of *RED<sub>exp</sub>*, one must be able to identify a *red<sub>exp</sub>* sensation when experiencing one. One might also maintain that conceptual mastery of *RED<sub>exp</sub>* requires that one have experienced *red<sub>exp</sub>* and/or be able to imagine a *red<sub>exp</sub>* patch. Mary has never had a *red<sub>exp</sub>* experience, is unable to imagine a *red<sub>exp</sub>* sensation, and cannot identify a *red<sub>exp</sub>* experience as an instance of her concept *RED<sub>exp</sub>*. Therefore, she does not have conceptual mastery of *RED<sub>exp</sub>*. As a result, Mary's inability to know propositions containing *RED<sub>exp</sub>* on the basis of the austere physical information shows nothing about necessitation. Only if such an epistemic gap occurs for an agent with mastery of *RED<sub>exp</sub>* can we begin to make claims about necessitation.

---

Footnote 15 continued

my own. The strategy involves identifying a philosophical argument or position that implicitly assumes that possession of a concept entails mastery. (This assumption is often masked by an ambiguity, between possession and mastery, in the terminology of "understanding/possessing a concept"). This assumption undermines the position. Once we restrict attention to cases in which the assumption holds, the argument fails to accomplish its goal. For example, I will argue (Sects. 9–11) that once we focus on cases in which Mary has conceptual mastery, she can know the target experiential propositions.

<sup>16</sup> For obvious reasons, the epistemic agent must have conceptual mastery of all concepts in the base propositions as well.

<sup>17</sup> The missing concept reply has appeared in the literature before (e.g. Hellie (2004), Loar (1990), Stoljar (2005), and Tye (2000)). To the best of my knowledge, the "conceptual mastery" reply has not.

## 7 The phenomenal concept strategy

We can delineate three different requirements on possession of an experiential/phenomenal concept.

- *Experiential requirement*: Possession of an experiential concept requires having had experiences that fall under the concept.
- *Imaginative requirement*: Possession of an experiential concept requires the ability to imagine instances of the phenomenal quality that falls under the concept.
- *Recognitional requirement*: Possession of an experiential concept requires the ability to recognize experiences that fall under the concept as falling under the concept.

Fans of the phenomenal concept strategy have endorsed versions of these requirements, and others still.<sup>18</sup> I deny them all; concept possession is too easy to obtain.

Historically, defenders of so-called “phenomenal concept strategies” have endorsed the experiential requirement on possession of a phenomenal concept.<sup>19</sup> Mary has not experienced  $\text{red}_{\text{exp}}$  and does not possess the concept  $\text{RED}_{\text{exp}}$ . Since she cannot consider the relevant propositions, her inability to know them demonstrates nothing about physicalism. This strategy has been popular, but less than effective.

Once the possibility of possessing a concept without mastery, potentially through deference to others, is recognized, the experiential requirement looks implausible. Blind people, who have never seen  $\text{red}_{\text{exp}}$ , can possess the concept  $\text{RED}_{\text{exp}}$ . They can truly think and say, “Tomatoes cause  $\text{red}_{\text{exp}}$  experiences.” Ball (2009) and Tye (2009) correctly press this objection against the experiential requirement and its use against the knowledge argument.<sup>20</sup> They point out that one can run a knowledge argument in which Mary does possess the relevant phenomenal concepts, but in which this possession will not help her learn the truths about experience (this is effectively the version of the argument I considered at the beginning of Sect. 6). Unfortunately, this move does not cut much ice. As argued in the previous section, the epistemic capabilities of someone who merely possesses the concept indicate nothing about necessitation, and hence cannot be brought to bear on the question of

<sup>18</sup> Papineau (1998, p. 5) endorses the experiential requirement, writing that anyone “who has never seen anything red cannot deploy a phenomenal concept of red visual experience.” Loar (1990) endorses the recognitional requirement. For a good discussion cf. Stoljar (2005).

<sup>19</sup> Papineau (1998) provides one example.

<sup>20</sup> Ball (2009) and Tye (2009) use the term ‘phenomenal concept’ for any concept possession of which requires having had experiences that fall under that concept. On this terminology, the experiential requirement on possession of a phenomenal concept is true by definition. Ball denies that there are any such concepts. But neither Ball nor Tye denies that there are experiential or phenomenal concepts in my sense—concepts tied to looking like *this* (where a red patch is demonstrated). Ball and Tye also use the term ‘phenomenal concept strategy’ for a response to the knowledge argument that endorses the experiential requirement on possession of experiential/phenomenal concepts. I use ‘phenomenal concept strategy’ more generally, applying it to any view that appeals to special features of experiential or phenomenal concepts to explain how necessitation from the physical to the experiential is compatible with a lack of epistemic implication.



physicalism. Epistemic considerations involving agents with less than conceptual mastery are simply a non-starter.

The experiential, imaginative, and recognitional requirements on concept possession have obvious analogs for conceptual mastery. I do not bother to state them. All three requirements are more plausible for conceptual mastery than for concept possession. One can't draw metaphysical conclusions from absence of epistemic implication unless the epistemic agents have conceptual mastery. Thus, as a reply to the knowledge argument, an experiential requirement on conceptual mastery of experiential/phenomenal concepts can do all the same work as an experiential requirement on possession. But the experiential requirement on conceptual mastery of  $RED_{exp}$  is not subject to the Ball/Tye-style criticism from deference. The original experiential requirement (on possession) looks bad because it's clear that Mary can possess  $RED_{exp}$  without ever having had a  $red_{exp}$  experience. On the other hand, it's not obvious that Mary can have conceptual mastery of  $RED_{exp}$  without having had such an experience. I advise fans of the original experiential requirement to transfer their allegiances to the experiential requirement on conceptual mastery. It can do all the same work in replying to the knowledge argument without incurring the obvious costs.<sup>21</sup> Furthermore, once we distinguish between concept possession and conceptual mastery, the conceptual mastery reply is a natural extension of the original phenomenal concept strategy.

I neither endorse nor deny the experiential or the imaginative requirement on conceptual mastery of experiential concepts. The experiential requirement seems counterexample-able by a "swamp case". Take a conceptual master of  $RED_{exp}$ . Imagine an atom-by-atom duplicate, generated by erratic random forces in the swamp mist. It is plausible that the duplicate will have conceptual mastery of  $RED_{exp}$  despite never having seen a red thing. Whether one can be a conceptual master of  $RED_{exp}$  without satisfying the imaginative requirement seems more questionable. However, one might be willing to attribute conceptual mastery to creatures who lacked certain cognitive imaginative capacities, but nonetheless had experienced  $red_{exp}$  and were very reliable in their  $red_{exp}$  judgements.

Perhaps the requirements can be softened by including a clause about "normal conditions"—i.e. conditions that allow for exceptions involving erratic swamp forces, genies, and fantastical neuro-surgery. Officially, I endorse only the

<sup>21</sup> Alter (manuscript) replies to Ball (2009) and Tye (2009) in a similar way, and offers the same advice. He also makes the important observation that the epistemic capabilities of mere possessors of a concept are not good indicators of metaphysical necessitation. However, this observation plays a far less central role in his discussion than in mine. Alter has his own version of the concept possession/conceptual mastery distinction. He uses the terminology of possessing a concept under a "deferential" or "non-deferential" mode of presentation. This terminology is misleading for two reasons. First, on the most natural extension of Fregean modes of presentation to our current framework, concepts *are* modes of presentation. Talk of modes of presentation *of* modes of presentation is misleading at best. Second, whether an individual *defers* to someone else, in the sense of accepting correction regarding the use of his or her concept, is only a defeasible guide to whether he or she has mere concept possession or mastery. One can use the term 'deferential possession' as a technical term for "possessing a concept without conceptual mastery, in a manner similar to Burge's arthritis-man does". But if deferential possession is *defined* in this way, then there is an open question about whether the fact that an individual defers, in the sense of accepting correction, implies anything about whether he "deferentially possesses" the concept in the defined sense.

recognitional requirement on conceptual mastery of experiential concepts.<sup>22</sup> However, even without the experiential requirement on conceptual mastery, it remains quite plausible that Mary does not have conceptual mastery of  $\text{RED}_{\text{exp}}$ . Her inability to know experiential propositions cannot be wielded against physicalism.

## 8 The informational assumption and the second horn

The second premise of the knowledge argument requires an inference of the following type: “ $\mathcal{P}$  does not implicate T, therefore  $\mathcal{P}$  does not necessitate T.” The contrapositive of this inference is: “ $\mathcal{P}$  necessitates T, therefore  $\mathcal{P}$  implicates T.” In this section, we turn our attention to this contrapositive. I lay out the *Informational Assumption*, which says that Mary can know every proposition necessitated by  $\mathcal{P}$ . The knowledge argument needs the informational assumption for two reasons. First, the informational assumption provides the link between epistemology and metaphysics on which the knowledge argument relies. Second, not making the informational assumption hands the physicalist an easy reply to the knowledge argument.

Suppose the informational assumption is false. Then there is some proposition R (perhaps the proposition that the rate of inflation in Australia is increasing), necessitated by  $\mathcal{P}$ , that Mary cannot know. Suppose also that premise (1) of the knowledge argument is true—despite complete physical knowledge, Mary can’t know that tomatoes cause  $\text{red}_{\text{exp}}$  sensations. The dualist wants to conclude that  $\mathcal{P}$  does not necessitate that tomatoes cause  $\text{red}_{\text{exp}}$  sensations. But the existence of exceptions to the informational assumption gives the physicalist an immediate reply. She’ll say that the proposition that tomatoes cause  $\text{red}_{\text{exp}}$  sensations is, like R, one among many propositions that are necessitated by  $\mathcal{P}$  yet not knowable by Mary. To cut off this reply the dualist needs the informational assumption—there are no propositions like R.

From here on, we make the informational assumption, and assume that Mary, using her knowledge of  $\mathcal{P}$ , can know every proposition necessitated by  $\mathcal{P}$ .<sup>23</sup> With

<sup>22</sup> Disclaimer for the counterexample mongers: I don’t endorse the *crazy* version of the recognitional requirement. A conceptual master of  $\text{BLUE}_{\text{exp}}$  need not be infallible in her  $\text{blue}_{\text{exp}}$  judgements. If she is drunk, temporarily mentally impaired, or some such thing, all bets are off.

<sup>23</sup> There may be some exceptions to the informational assumption. For example, suppose dualism is true and the experiential truths are not necessitated by the austere physical state of the world described in  $\mathcal{P}$ . Let ‘Bob’ be a name whose reference is fixed by the description ‘the actual greenest thing in the universe’. Suppose Bob is, in fact, a meadow in Vermont. It is plausible that the proposition  $\text{BOB HAS SURFACE AREA OF MORE THAN } 10 \text{ M.}^2$  is necessitated by  $\mathcal{P}$  despite the fact that someone who knows  $\mathcal{P}$  will not (if dualism is true) be in a position to know it (because he or she will not be in a position to know the experiential facts about  $\text{green}_{\text{exp}}$ ). We seem to have a counterexample to the informational assumption. The lesson is that the knowledge argument is compatible with the existence of *some* propositions that are both necessitated by  $\mathcal{P}$  and not knowable on the basis of  $\mathcal{P}$ . Advocates of the knowledge argument must hope that there are not too many exceptions to the informational assumption. Every exception opens up space for the physicalist to reply to the knowledge argument with the claim that the proposition that tomatoes cause  $\text{red}_{\text{exp}}$  sensations is also an exception, and hence Mary’s inability to know it does not demonstrate a lack of necessitation. Thankfully, the exceptions look to be of a special sort. Many of them will contain rigidified descriptions involving phenomena not necessitated by the physical (e.g. dualistic

the help of the informational assumption, we precisify the basic dilemma of this paper.

**The Dilemma:** Either Mary does not have conceptual mastery of experiential concepts, in which case we cannot draw any metaphysical conclusions on the basis of what she can and cannot know, or Mary does have conceptual mastery, and, with the help of knowledge allotted to her by the informational assumption, she will be able to know all the relevant experiential propositions.

The dilemma is related to my claim that it is difficult to find a notion of epistemic implication that satisfies both premises of the knowledge argument schema of Sect. 5. On the first horn of the dilemma, Mary lacks conceptual mastery, and we cannot draw metaphysical conclusions from her case. On any notion of implication that does not require epistemic agents to have conceptual mastery, premise (2) of the schema is false. On the second horn, Mary has conceptual mastery. But this mastery, in combination with the knowledge allotted to Mary by the informational assumption, will generate an epistemic route to knowledge of experiential propositions. On this horn, premise (1) of the schema is false—Mary can know.

I have already argued for the first horn of the dilemma. In the next three sections I argue for the second horn. I demonstrate exactly how Mary's conceptual mastery of, e.g.  $RED_{exp}$ , in combination with the informational assumption, yields an epistemic route to knowledge that, e.g., tomatoes cause  $red_{exp}$  sensations.

Advocates of the knowledge argument grant that the informational assumption gives Mary knowledge of every proposition necessitated by  $\mathcal{P}$ . But the dualist and physicalist disagree about which propositions these are. The physicalist claims that experiential propositions are necessitated; the dualist denies. In order to wield the informational assumption against the knowledge argument, we must apply it only to propositions the dualist admits  $\mathcal{P}$  necessitates. I introduce a heuristic for determining which propositions these are. To determine whether the dualist grants that  $\mathcal{P}$  necessitates a proposition, check whether the proposition is true at both the actual and zombie worlds. The *zombie world* is an austere physical duplicate of the actual world that contains no experiential properties. Zombies walk the walk and talk the talk (they even do zombie philosophy), but there is nothing "it is like" to be a zombie.<sup>24</sup>

Our basic strategy is to play along with the dualist, using the zombie world to get him to agree that certain propositions are necessitated by  $\mathcal{P}$ , and thus knowable by Mary. We then use Mary's knowledge to refute the claim that Mary cannot know that tomatoes cause  $red_{exp}$  sensations.<sup>25</sup> Consider some applications of the heuristic. Rocks exist at both the actual and zombie worlds; the dualist grants that  $\mathcal{P}$  necessitates that rocks exist. By the informational assumption, Mary knows that

---

Footnote 23 continued

experiential properties). In this paper, the informational assumption will be brought to bear only on propositions that are not of this special type. If the reader thought that many or most terms (or concepts) were rigidified versions of phenomenal descriptions, then the informational assumption would not be of much use. I assume this is not the case.

<sup>24</sup> Chalmers (1996) deserves credit both for the invention of philosophical zombies and for bringing them into the contemporary discussion.

<sup>25</sup> Use of the zombie heuristic does not require that the zombie world be metaphysically possible.

rocks exist. More interestingly, the dualist acknowledges that people in both the zombie and actual worlds enter neural state N whenever they see a ripe tomato. Thus  $\mathcal{P}$  necessitates the proposition that a person enters neural state N whenever they see a ripe tomato; Mary can know it. On the other hand, the dualist denies that, at both the actual and zombie worlds, people experience  $\text{red}_{\text{exp}}$  whenever they enter neural state N. The zombie heuristic tells us we cannot assume, via the informational assumption, that Mary knows this.<sup>26</sup>

## 9 Re-captured Mary

In order to draw metaphysical conclusions from what Mary can and cannot know, we must insist that Mary has conceptual mastery of experiential concepts, including  $\text{RED}_{\text{exp}}$ . In the knowledge argument as originally described, Mary did not have such mastery. But there is an obvious suggestion for how to proceed. If Mary did have conceptual mastery of  $\text{RED}_{\text{exp}}$ , would she be able to know the troublesome experiential propositions?

Mary can acquire mastery of experiential color concepts if we remove her from the black and white room and expose her to all the experiences that are necessary for her to acquire conceptual mastery (e.g. experiences of red tomatoes, blue sky, green grass, yellow lemons, etc.). Then bring Mary back to the black and white room and give her a pill that makes her forget her expedition outside. Mary has now experienced  $\text{red}_{\text{exp}}$  sensations, has the ability to imagine  $\text{red}_{\text{exp}}$  patches, and can identify them as  $\text{red}_{\text{exp}}$  (likewise for the other colors). But she does not (yet) know that tomatoes cause  $\text{red}_{\text{exp}}$  sensations. *Prima facie*, it seems Mary's conceptual mastery of experiential concepts will not help her come to know experiential propositions. She's got the color concepts,  $\text{RED}_{\text{exp}}$ ,  $\text{BLUE}_{\text{exp}}$ , and  $\text{GREEN}_{\text{exp}}$ , and she knows what experiences of each type are like. But she does not know how to map these colors on to the world. She does not know which of these colors she can so vividly imagine is caused by tomatoes and which by the sky. Thus, there may be some propositions that Mary cannot know, despite her complete austere physical knowledge and her conceptual mastery. If so, perhaps the knowledge argument can be salvaged.<sup>27</sup>

Which proposition is it, exactly, that Mary cannot know? Is it the proposition that tomatoes cause  $\text{red}_{\text{exp}}$  sensations? Mary can, with surprising ease, come to know this proposition. She does not even need conceptual mastery of  $\text{RED}_{\text{exp}}$ . Mary knows that an overwhelming majority of English speakers who have seen tomatoes testify that

<sup>26</sup> The zombie heuristic is fallible. A proposition might be true at both the actual and zombie worlds, but false at some other world where  $\mathcal{P}$  is true. Maybe there are exotic worlds at which  $\mathcal{P}$  is true but rocks do not exist. This is unlikely, but a proposition's truth at the actual and zombie worlds does not guarantee its necessitation by  $\mathcal{P}$ . Here, we apply the informational assumption only to propositions that are plausibly necessitated by  $\mathcal{P}$  and granted by dualists as such.

<sup>27</sup> The use of "re-captured Mary", or a version of Mary who has had the relevant experiences, is not new. Versions appear in Lewis (1988), Nida-Rumelin (1996), and Stoljar (2005). I know of no use, other than my own, of re-captured Mary as a method of enabling Mary to acquire conceptual mastery, rather than possession, of experiential concepts.

they cause  $\text{red}_{\text{exp}}$  sensations.<sup>28</sup> This testimony seems perfectly sufficient for Mary to learn that tomatoes cause  $\text{red}_{\text{exp}}$  sensations.

Perhaps the proposition that Mary cannot know is expressed by ‘ $\text{red}_{\text{exp}}$  sensations feel like this’ (or something similar). The knowledge Mary allegedly lacks is sometimes stated this way. However, it is unclear what the ‘this’ stands for. If ‘this’ is a pure demonstrative, then it is no surprise that Mary cannot come to know that “ $\text{red}_{\text{exp}}$  sensations feel like this”. This proposition can only be known by someone in the presence of a  $\text{red}_{\text{exp}}$  experience. Just as one cannot come to know “this is a chair” unless one is in the presence of a chair, one cannot come to know that “ $\text{red}_{\text{exp}}$  sensations feel like this” unless one is having a  $\text{red}_{\text{exp}}$  experience. In the black and white room, Mary is not having an occurrent  $\text{red}_{\text{exp}}$  experience, so she can’t consider the relevant proposition, let alone know it. This shows nothing about necessitation.

Mary might imagine a  $\text{red}_{\text{exp}}$  patch, demonstrate it, and consider the proposition expressed by ‘ $\text{red}_{\text{exp}}$  sensations feel like this’, where ‘this’ refers to the qualitative feel of the  $\text{red}_{\text{exp}}$  patch. But Mary, if she is a master of the concept  $\text{RED}_{\text{exp}}$ , can know this proposition. By the recognitional requirement, mastery of  $\text{RED}_{\text{exp}}$  requires the ability to identify such feels as  $\text{red}_{\text{exp}}$ . If Mary is a master of  $\text{RED}_{\text{exp}}$ , she knows that the proposition expressed by ‘ $\text{red}_{\text{exp}}$  sensations feel like this’ is true. If she does not know that this proposition is true, then she can’t identify  $\text{red}_{\text{exp}}$  sensations, and ipso facto she lacks conceptual mastery of  $\text{RED}_{\text{exp}}$ . Either way, the knowledge argument is answered.

## 10 Wow: re-captured Mary redux

Mary’s epistemic route to knowledge that tomatoes cause  $\text{red}_{\text{exp}}$  sensations, sketched above, may not satisfy the reader. Reliance on the testimony of others seems suspect, and irrelevant to the main thrust of Jackson’s original argument. We can press the point by modifying the case.

As before, let Mary experience color sensations during a field trip outside the black and white room. Drug her so she cannot remember which of the colors she can now vividly imagine is the color of tomatoes, and which the color of the sky. Mary retains the ability to imagine  $\text{red}_{\text{exp}}$  patches. She introduces a term, ‘wow’, and an experiential concept, wow, which applies to experiences with the qualitative feel of  $\text{red}_{\text{exp}}$  patches.<sup>29</sup> Mary plausibly has conceptual mastery of wow. She has experienced wow sensations (otherwise known as  $\text{red}_{\text{exp}}$  sensations), has the ability to imagine wow patches, and can identify them as instances of wow (likewise for the other colors). But Mary does not know which of tomatoes, the sky, or grass, is

<sup>28</sup> Most speakers of English don’t distinguish between the concepts  $\text{RED}$  and  $\text{RED}_{\text{exp}}$ . But this should not impede Mary’s ability to learn from testimony that tomatoes cause  $\text{red}_{\text{exp}}$  sensations. There’s an obvious principle linking the two: red things cause  $\text{red}_{\text{exp}}$  sensations.

<sup>29</sup> I use the term ‘wow’ in deference to Perry (2001), who has a similar example, and introduces the word.

wow-colored. This time, Mary cannot rely on others' testimony about 'wow'—there is no such testimony. If Mary cannot know that tomatoes cause wow sensations, the knowledge argument looks to be in decent shape—there are still some propositions Mary cannot know.

I argue that Mary can know that tomatoes cause wow sensations. The amnesia pill ensures that Mary does not remember her expedition outside the black and white room from the first-person perspective. But because Mary knows  $\mathcal{P}$  and every proposition necessitated by it, she knows all the details of her expedition from a third-person perspective. She knows that she entered neural state N whenever she saw a ripe tomato, strawberry, or fire engine, during that expedition. She knows that whenever she imagines a wow patch while sitting in the black and white room, she also enters neural state N. She'll learn that N is the neural correlate of wow experiences, and that tomatoes cause the instantiation of N. This information will be sufficient for her to learn that tomatoes cause wow sensations.

## 11 Lonely Mary and the pincer argument

The reader may be worried about the preceding arguments. In Sect. 9, I argued that, by relying on the testimony of others, Mary can know that tomatoes cause  $\text{red}_{\text{exp}}$  sensations. In Sect. 10, I argued that Mary can know that tomatoes cause wow sensations (otherwise known as  $\text{red}_{\text{exp}}$  sensations) by combining (3rd person) facts about her expedition outside the black and white room with her capacity to imagine wow-colored patches. Staunch defenders of the knowledge argument will deny that having experienced  $\text{red}_{\text{exp}}$  is a necessary condition on conceptual mastery of  $\text{RED}_{\text{exp}}$ , and also deny that the ability to imagine a  $\text{red}_{\text{exp}}$  patch is required. To satisfy the doubters, I argue that even in a Mary case in which (a) no one else ever exists, (b) Mary never leaves the black and white room, (c) conceptual mastery of an experiential concept does not require having had experiences of the appropriate type, and (d) conceptual mastery does not require an ability to imagine the experiential quality, Mary can *still* know the target experiential propositions, e.g., that tomatoes cause  $\text{red}_{\text{exp}}$  sensations.

Let Mary be the only person ever to have existed. She has complete knowledge of  $\mathcal{P}$ . But this time suppose that Mary has never left the black and white room, never experienced  $\text{red}_{\text{exp}}$ , and is (for whatever reason) unable to imagine a  $\text{red}_{\text{exp}}$  patch. But nonetheless she has conceptual mastery of  $\text{RED}_{\text{exp}}$ , and (by the recognitional requirement) would identify a  $\text{red}_{\text{exp}}$  sensation, were she to have one, as an instance of  $\text{RED}_{\text{exp}}$ .

The *pincer argument* goes as follows:

- (1) Mary can know that tomatoes cause her to think the thought THIS SENSATION IS R.
- (2) Mary can know that the semantic value of the 'R' part of this thought is the concept  $\text{RED}_{\text{exp}}$ .
- (3) If Mary can know that tomatoes cause her to think THIS SENSATION IS R, and that 'R' means  $\text{RED}_{\text{exp}}$ , then she can know that tomatoes cause her to think THIS SENSATION IS  $\text{RED}_{\text{exp}}$ .

- (4) If Mary can know that tomatoes cause her to think THIS SENSATION IS RED<sub>exp</sub>, then she can know that tomatoes cause red<sub>exp</sub> sensations.<sup>30</sup>

- (5) Therefore, Mary can know that tomatoes cause red<sub>exp</sub> sensations.

I briefly sketch the support for each premise here; then I go into detail. The informational assumption ensures that Mary knows lot of facts about brain states and about how those brain states relate to representational states, especially to the syntax of those representational states. These facts will get Mary to the knowledge attributed to her in premise (1). Mary will integrate knowledge about her brain state at a time and knowledge about what she was thinking at that time to acquire the knowledge premise (2) attributes to her. Premise (3) is obvious. Mary can know that she is not in a deviant case, which is all that is needed to acquire the knowledge attributed to her in premise (4).

Since Mary has conceptual mastery of RED<sub>exp</sub>, the recognitional requirement entails that Mary is disposed to think THIS SENSATION IS RED<sub>exp</sub> (not THIS SENSATION IS BLUE<sub>exp</sub>) whenever she sees a ripe tomato. Mary learns that tomatoes cause her to think THIS SENSATION IS RED<sub>exp</sub> by combining two bits of information. First, she'll learn that tomatoes cause her to think THIS SENSATION IS R (premise (1)). Second, she'll learn that the semantic value of the 'R' portion of this thought is her concept RED<sub>exp</sub> (premise (2)).<sup>31</sup> Premise (3) says Mary can combine these two pieces of information in the obvious way.

The proposition that tomatoes cause Mary to think THIS SENSATION IS R is necessitated by  $\mathcal{P}$ . By the informational assumption, Mary knows it. Evidence for this necessitation comes from considering the zombie world. Zombie Mary also thinks THIS SENSATION IS R. The difference between Mary and zombie Mary is that Mary's mental type 'R' has as its semantic value RED<sub>exp</sub>, whereas zombie Mary's 'R' means RED<sub>zombie</sub>. Actual speakers and zombies both mean THIS, IS, and SENSATION by their uses of 'this', 'is', and 'sensation', respectively. The dualist should grant that these semantic facts are necessitated by  $\mathcal{P}$ , and grant that  $\mathcal{P}$  necessitates that tomatoes cause Mary to think THIS SENSATION IS R.<sup>32</sup> By the informational assumption, Mary can know that tomatoes cause her to think THIS SENSATION IS R. Premise (1) is true.

To complete her epistemic route to knowledge that tomatoes cause Mary to think THIS SENSATION IS RED<sub>exp</sub>, Mary needs only to learn that the 'R' component means RED<sub>exp</sub>. The informational assumption ensures Mary knows the state of her own brain at every moment in history, both past and future. She also knows how her brain state relates to, and realizes, certain features of her representational or intentional state. For example, the syntactic structure of Mary's representational

<sup>30</sup> The causation here must be of the appropriate type. For example, if tomatoes miraculously cause Mary to think THIS SENSATION IS RED<sub>exp</sub> whenever they strike her forehead, there is no reason to think that tomatoes cause red<sub>exp</sub> sensations. I leave the qualification about causation implicit.

<sup>31</sup> When I speak of 'semantic value' here, I mean (something like) Fregean sense, or meaning, not extension. On this terminology, the semantic values of 'this' and of 'Mary' differ, even on an occasion where a token of 'this' demonstrates Mary.

<sup>32</sup> In Sect. 12, I address how the argument fares if the dualist denies that some of these facts are necessitated.



state is plausibly shared by zombie Mary, and necessitated by  $\mathcal{P}$ . Therefore, Mary knows the syntactic structure of her representational state at every moment in history, and she knows which aspects of her brain realize which aspects of that syntactic state. It follows that Mary knows exactly what brain state-type realizes the ‘R’ token-type of her representational state. Mary knows exactly at which moments she has thought, and will think, an ‘R’ token, because she knows when her brain is in the state that realizes the ‘R’ token.

Mary needs to figure out the semantic value of her ‘R’ tokens. To do this, she’ll actively consider, at some time  $t$ , the thought  $\text{ALL RED}_{\text{exp}} \text{ THINGS ARE RED}_{\text{exp}}$ . By the informational assumption, Mary already knew that, at time  $t$ , two occurrences of the ‘R’-element would occur in her brain’s “actively considering box” (an analog of the belief box). Once Mary knows that at time  $t$  she thought  $\text{ALL RED}_{\text{exp}} \text{ THINGS ARE RED}_{\text{exp}}$  (and not  $\text{THIS MANGO IS A SUCCULENT MANGO}$ ), she’ll learn that ‘R’ means  $\text{RED}_{\text{exp}}$ , and not  $\text{MANGO}$ , or  $\text{BLUE}_{\text{exp}}$ . Therefore, premise (2) is true.

Premise (3) is obvious. It says that if Mary can know both that tomatoes cause her to think  $\text{THIS SENSATION IS R}$  and that ‘R’ means  $\text{RED}_{\text{exp}}$ , then she can know that tomatoes cause her to think  $\text{THIS SENSATION IS RED}_{\text{exp}}$ .

Premise (4) says that Mary can infer from the fact that tomatoes cause her to think  $\text{THIS SENSATION IS RED}_{\text{exp}}$  to the fact that tomatoes actually cause  $\text{red}_{\text{exp}}$  sensations. To do this, Mary needs to check that she is not being misled in the particular case of tomatoes. For example, if Mary applied the concept  $\text{RED}_{\text{exp}}$ , for the most part, to all and only  $\text{red}_{\text{exp}}$  experiences, but misapplied to tomato-caused  $\text{blue}_{\text{exp}}$  sensations (perhaps because tomatoes trigger her allergies), then tomatoes would cause Mary to think  $\text{THIS SENSATION IS RED}_{\text{exp}}$  without tomatoes actually causing  $\text{red}_{\text{exp}}$  sensations. Mary knows she is not in such a deviant case. She knows that tomatoes, strawberries, and fire engines all emit light of wavelength 700 nanometers, have dramatically similar effects on the color processing component of her visual system, and generate the same sensational neural correlate. She knows tomatoes do not trigger an allergic reaction that causes her system to misfire and label her experiences in a non-standard way. Premise (4) is true.

Premises (1)–(4) of the pincer argument are true, and the argument is valid. The conclusion follows: Mary can know that tomatoes cause  $\text{red}_{\text{exp}}$  sensations. By combining (i) knowledge “from below” about brain states and the syntax of the representational states those brain states realize (knowledge allotted to Mary by the informational assumption) and (ii) knowledge “from above” about what thought ( $\text{ALL RED}_{\text{exp}} \text{ THINGS ARE RED}_{\text{exp}}$ ) she was thinking at a given time, Mary can know that tomatoes cause her to think  $\text{THIS SENSATION IS RED}_{\text{exp}}$ .<sup>33</sup> From there, it’s only a short step to knowledge that tomatoes cause  $\text{red}_{\text{exp}}$  sensations.

<sup>33</sup> The “pincer argument” gets its name from Mary’s two-pronged approach to knowledge that tomatoes cause her to think  $\text{THIS SENSATION IS RED}_{\text{exp}}$ . In military strategy, a simultaneous attack from two sides is called a ‘pincer movement’ or ‘double envelopment’.



## 12 Objections

In this section, I consider six objections. The first two are objections to claims about intentionality used in the pincer argument, the third is an objection to the recognitional requirement on conceptual mastery of RED<sub>exp</sub>, the fourth and fifth are objections to my presentation of the knowledge argument, and the sixth is an objection to my story about the connection between conceptual mastery and the modal rationalist inference.

One objection to the pincer argument says that the intentional facts I claim Mary knows are not necessitated by  $\mathcal{P}$ , and thus the informational assumption does not give Mary knowledge of them—premise (1) is false. The important question here is “Are the intentional propositions needed by Mary to learn that tomatoes cause red<sub>exp</sub> sensations necessitated by  $\mathcal{P}$ ?” If the answer is ‘yes’, then these propositions are, by the informational assumption, known by Mary, and the objection is answered. The objector, a dualist, claims the answer is ‘no’.

One view that supports this objection claims that all intentional facts depend on phenomenological/experiential facts—phenomenology is prior to intentionality. Something like this view is suggested in Horgan and Tienson (2002), Graham et al. (2007), and Strawson (1996). If this is correct, and zombies are possible, then there is no intentionality at the zombie world, no representational states are necessitated by  $\mathcal{P}$ , and no intentional facts are knowable by Mary via the informational assumption. However, the “phenomenology first” view couples badly with dualism. If you believe that zombies are possible, you should believe that at least some intentional facts are independent of phenomenal experience. Zombies appear to utter sentences and think thoughts. At least some of these sentences and thoughts are true, others false. If so, then at least some intentionality floats free of experiential phenomenology. For example, it is very plausible that the syntactic form of sentences, and of thoughts, does not depend on phenomenology, and thus does not vary between Mary and zombie Mary. Furthermore, it is plausible that the meaning of many terms, such as ‘is’, ‘that’, and ‘basketball’, does not shift between the actual and zombie worlds. If these two claims are correct, then the pincer argument can fend off the objection. That argument required only that  $\mathcal{P}$  necessitates the proposition that tomatoes cause Mary to think THIS SENSATION IS R. The dualist should admit that enough intentionality is independent of phenomenology to make this proposition true at the zombie world. If so, the proposition is plausibly necessitated by  $\mathcal{P}$  and knowable by Mary, as premise (1) claims.

The second objection, closely related to the first, claims that the meaning of ‘sensation’ varies between the actual and zombie worlds. Since zombies have no phenomenal experiences, they mean something different by ‘sensation’. But I claimed that ‘sensation’ has the same meaning in our mouths that it does in zombie mouths. I reply that the claim that ‘sensation’ has the same meaning at the actual and zombie worlds is inessential to the pincer argument. Even if the meaning shifts, Mary can learn the meaning of ‘sensation’ using the same trick she used to learn the meaning of ‘R’. By the informational assumption, she’ll know that tomatoes cause her to think THIS S IS R. She can then use the method previously described (of

thinking thoughts containing the concept  $RED_{exp}$ , or SENSATION), to deduce that ‘S’ means SENSATION.

This reply offers a lesson and bolsters the reply to the first objection. The pincer argument requires only that facts about the syntactic structure of Mary’s representational states are necessitated by  $\mathcal{P}$  (and thus knowable by Mary according to the informational assumption). Once Mary knows she is thinking a thought with a given syntactic form, and is able to re-identify mental language tokens as being of the same type, she can use the familiar trick to determine the semantic values of those tokens. It is tremendously plausible that the syntax of an agent’s language (both spoken and mental) is necessitated by  $\mathcal{P}$  and shared by Mary and her zombie twin.

The pincer argument relies on the informational assumption and Mary’s conceptual mastery of  $RED_{exp}$ . Only the recognitional requirement on conceptual mastery—not the experiential or imaginative requirement—is needed. Using only the recognitional requirement gives the argument wider appeal. Deniers of the experiential and imaginative requirements can be convinced. But the true skeptics will remain doubtful. Why think that conceptual mastery of the experiential concept  $RED_{exp}$  (or *wow*) requires so much? Why think that recognitional capacities are a requirement on conceptual mastery of  $RED_{exp}$ , or on experiential concepts in general? I do not have a knock-down argument against the dedicated skeptic. I ask the reader to consider a case, and consult his or her intuition.

Conceptual mastery is the type of understanding had by the proverbial “experts” to whom users who merely possess the concept, but do not have mastery, defer. The thought contents of mere possessors are determined by those who have full-fledged conceptual mastery. Imagine a case in which conceptual mastery of  $RED_{exp}$  does not require the ability to recognize a  $red_{exp}$  experience as  $RED_{exp}$ . Suppose we are all blind, but use the concept  $RED_{exp}$ , and defer in our use to the sighted master, whom we call ‘Sensei’. The reference, content, and Fregean sense of our  $RED_{exp}$  concept is determined by Sensei’s. Sensei, when presented with three color sensations, one  $red_{exp}$ , one  $blue_{exp}$ , and one  $green_{exp}$ , and asked, “which of these sensations is  $RED_{exp}$ ?” will throw up his hands and say, “I have no idea!” Given that Sensei is the ultimate authority here (he does not defer to further experts), it’s unclear what makes the concept  $RED_{exp}$  have  $red_{exp}$  rather than  $blue_{exp}$  sensations in its extension. One way to ensure that  $RED_{exp}$  refers to  $red_{exp}$  sensations is to have the expert to whom everyone defers identify  $red_{exp}$ , and not  $blue_{exp}$ , sensations as falling under  $RED_{exp}$ . Conceptual mastery is that which is had by the proverbial experts. The experts regarding  $RED_{exp}$  must be able to recognize  $red_{exp}$  sensations as such; otherwise they aren’t experts, and don’t have conceptual mastery.

An objector might disagree with my interpretation of the knowledge argument. He or she might think that I’ve targeted the wrong propositions and claim that there is some other proposition that Mary does not know. I disagree: there is no such proposition. By the informational assumption, Mary has a tremendous amount of information at her disposal. This information will allow her to know any proposition you like, often by means similar to those I have outlined.

An objector might instead maintain that the knowledge Mary lacks is not propositional. I have no beef with this claim. I maintain only that there is no

propositional knowledge that Mary lacks.<sup>34</sup> There are two reasons why propositional knowledge is what matters here. First, if the knowledge Mary lacks is not propositional, it is difficult to see how her ignorance entails the falsity of the physicalist's necessitation claim. Necessitation is a relation between propositions (or perhaps between states of affairs, which are easily translated into the language of propositions). Second, the claim that the knowledge Mary lacks is not propositional has historically been used as a physicalist *defense* against the knowledge argument, not as a method of resurrecting the argument after losing the initial battle by admitting that Mary lacks no propositional knowledge. Two examples of this physicalist strategy include the ability hypothesis of Lewis (1988) and Nemirow (1990) and the acquaintance hypothesis of Conee (1994).

Lastly, an objector might not buy my arguments regarding concept possession and conceptual mastery. He or she might think that one can infer from an inability to know, even for an agent who lacks the relevant concepts (or who lacks full-fledged conceptual mastery), to an absence of necessitation. I think this line of thought grossly mistaken; I have already explained why.

### 13 Taking stock

One upshot here is that Mary, sitting in her black and white room, bestowed with conceptual mastery of experiential concepts such as RED<sub>exp</sub>, BLUE<sub>exp</sub>, and TICKLISH<sub>exp</sub>, *can* know every true proposition involving those concepts. This is a tremendously surprising result. It is worth noting that the result does not depend on the claim that physicalism is true. Even if dualism is true, Mary will *still* be able to know the experiential propositions.

My final position is very different than traditional phenomenal concept strategies, and I do not think my view is deserving of the name or the lineage. All phenomenal concept strategies accept that (a) Mary cannot know certain experiential truths on the basis of the physical truths and (b) the physical truths necessitate these experiential truths nonetheless. Special aspects of “phenomenal concepts” are then wheeled in to explain how (a) and (b) could both obtain. Phenomenal concepts explain how there could be necessitation from the physical to the experiential without epistemic implication. My strategy is very different. I claim that once we give Mary the conceptual capacities necessary to draw any metaphysical conclusions, Mary *will* be in a position to know the experiential truths. Thus I deny the datum (a) with which phenomenal concept theorists begin.

The most popular response to the knowledge argument is to deny the modal rationalist claim that one can draw metaphysical conclusions about necessitation from epistemological considerations. Phenomenal concepts supplement this response. Special features of phenomenal concepts are wheeled in to explain how there can be “opaque” necessitation—necessitation without epistemic implication.

<sup>34</sup> More accurately, I claim that there is no propositional knowledge Mary lacks in a way that challenges physicalism. There is no knowledge that Mary lacks for reasons other than lack of concept possession and/or conceptual mastery.

My response is much different. I'm ready to spot the advocate of the knowledge argument his modal rationalism. Let's assume that one *can* draw metaphysical conclusions on the basis of some epistemological notion.<sup>35</sup> Even then, I argue, we can preserve our physicalism in the face of the knowledge argument. This gives my account two dialectical advantage over most extant responses to the knowledge argument.<sup>36</sup>

First, a modal rationalist can endorse my reply, and not most others. Second, since most responses deny the modal rationalist connection between metaphysics and epistemology, the debate quickly shifts into deep issues in the philosophy of modality. With this shift comes the risk of devolution into a clash of basic intuitions about links between modality, knowability, a priority, and conceptual truth. I avoid the devolution by granting the modal rationalist his connections between epistemology and metaphysics, while still offering the physicalist an adequate reply to the knowledge argument.

Finally, I take pain not to overstate my position. Some philosophers, most notably Chalmers (1996) and Jackson (1998), interpret the knowledge argument as an attempted demonstration that there is no *a priori* implication from  $\mathcal{P}$  to the experiential truths. Have I demonstrated that there is such an *a priori* implication—at least for a subject who has conceptual mastery? This question is difficult. The answer depends on the relationship between *a priori* knowledge and introspection, and whether certain types of introspection can yield *a priori* justification.

In the pincer argument, Mary used a four step procedure to figure out the  $RED_{exp}$  facts. First, by the informational assumption, Mary knew that tomatoes, fire engines, and the like cause her to token 'R' representations. Second, she needed to determine that Rs are  $RED_{exp}$  representations and not  $BLUE_{exp}$  representations. She did this by actively thinking a thought involving the  $RED_{exp}$  concept. At the moment she does so her brain realizes a R token. From this she learns that R tokens are tokens of  $RED_{exp}$  (not of  $BLUE_{exp}$ ). Importantly, Mary had to actively consider a thought involving  $RED_{exp}$  at time *t* and use her knowledge that she considered that thought, and not another, at *t*. This move is certainly introspective.

I cannot address here the difficult questions surrounding the relationship between *a priori* justification and introspection. One move available to Chalmers, Jackson, and others, is to *define* a notion of *a priori* justification that does not allow introspection. They can then use Mary as evidence that there is no *a priori*<sub>no introspection</sub> implication from the austere physical facts  $\mathcal{P}$  to the experiential facts, and attempt to draw metaphysical conclusions. I admit that I have not shown that there is an *a priori*<sub>no introspection</sub> route from  $\mathcal{P}$  to the experiential facts, even if the agent has conceptual mastery. However, this stipulative move does not resolve the debate. The knowledge argument requires the existence of a link between epistemic

<sup>35</sup> Of course, I'm not willing to spot the advocate of the knowledge argument any inference he likes. Section 6 demonstrates that, if the modal rationalist claim that metaphysical conclusions follow from epistemology is to have any chance of being true, we must insist that the epistemic agents have conceptual mastery. The rationalist position I'm willing to play along with must acknowledge at least this much.

<sup>36</sup> Thanks to Ned Block for encouraging me to advertise this selling point and stress the differences between my account and those already in the literature.

implication and metaphysical necessitation. The a priori version of the knowledge argument requires an important (potentially constitutive) link between a priori implication and metaphysical necessitation. Defenders of the a priori knowledge argument can stipulate whatever sense of “a priori” they like. But they are not free to stipulate that it is a priori<sub>no introspection</sub>, rather than some *other* conception of the a priori, that has the vital link to metaphysical necessity. This latter claim requires argument.

For now, I leave it open (i) whether there is any notion of a priori implication that allows the drawing of conclusions about necessitation and (ii) if there is such a notion, whether it does or does not allow introspection. If it does not, then a strict a priori version of the knowledge argument might survive my criticisms. In this case, I will have shown only that there is an *armchair* implication from  $\mathcal{P}$  to the experiential facts. This result is itself significant. If the relevant notion of the a priori does allow introspection of the relevant type, then I take my arguments to demonstrate that there is an a priori implication from the complete physical facts  $\mathcal{P}$  to the experiential facts, and the knowledge argument is sunk.

## 14 Conclusion

The knowledge argument has a tough row to hoe. It requires a perfect storm of epistemological and metaphysical conditions. It requires that Mary have the experiences and capacities necessary to possess and master experiential concepts. And the argument requires that Mary simultaneously be unable to know experiential propositions and, by the informational assumption, able to know every proposition necessitated by the complete austere physical truth  $\mathcal{P}$ . That’s a lot to require.

The informational assumption gives Mary a lot to work with—information about the neural correlates of sensations, about what every human being has said throughout the history of time, about the relationship between brain states and intentional or representational states, and a tremendous deal more to boot. With such a plethora of information, it is not difficult to figure out the experiential truths themselves. Even in variations on the knowledge argument in which no one speaks the word ‘red’, or in which Mary is the only person ever to have existed, Mary can use the knowledge allotted to her by the informational assumption to figure out the experiential truths.

We could continue doing variations on the argument, on Mary’s situation, and on what proposition cannot be known, until the cows come home, but these basic constraints remain. We cannot draw any metaphysical conclusions unless Mary has conceptual mastery of experiential concepts. But once she has conceptual mastery, Mary will be able to know all the experiential propositions she likes. She’ll know about the redness of a ripe tomato, the greenness of a grassy field, and the brilliant blue of a clear sky on a cold winter day. The knowledge argument is a beautiful and provocative piece of philosophy. Jackson sketches a very promising line of anti-physicalist attack. But when it comes to the details, it is surprisingly difficult to fill in the sketch, and make good on Mary’s metaphysical potential.

**Acknowledgements** Thanks to Torin Alter, Ned Block, Bill Brewer, David Chalmers, Mark Greenberg, Brian Rabern, Jonathan Simon, the editors at *Philosophical Studies*, and audiences at New York University and the University of Warwick for helpful comments and criticism.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Non-commercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Alter, T. (manuscript). Social externalism and the knowledge argument.
- Ball, D. (2009). There are no phenomenal concepts. *Mind*, 118, 935–962.
- Bealer, G. (2002). Modal epistemology and the rationalist renaissance. In J. Hawthorne & T. S. Gendler (Eds.), *Conceivability and possibility* (pp. 71–125). Oxford: Oxford University Press.
- Burge, T. (1979). Individualism and the mental. *Midwest Studies in Philosophy*, IV, 73–122.
- Byrne, A. (2002). Something about Mary. *Grazer Philosophische Studien*, 63, 123–140. (Essays in honour of the philosophy of Terence Horgan).
- Byrne, A. (2006). Review of “There’s Something About Mary”. *Notre Dame Philosophical Reviews*. <http://ndpr.nd.edu/review.cfm?id=5561>.
- Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.
- Chalmers, D. (2003). Consciousness and its place in nature. In S. Stich & T. Warfield (Eds.), *The Blackwell guide to philosophy of mind* (pp. 102–142). Malden, MA: Blackwell.
- Chalmers, D. (forthcoming). *Constructing the world*. Oxford.
- Chalmers, D., & Jackson, F. (2001). Conceptual analysis and reductive explanation. *Philosophical Review*, 110, 315–361.
- Conee, E. (1994). Phenomenal knowledge. *Australasian Journal of Philosophy*, 72, 136–150.
- Graham, G., Horgan, T., & Tienson, J. (2007). Consciousness and intentionality. In M. Velmans & S. Schneider (Eds.), *The Blackwell companion to consciousness* (pp. 468–484). Malden, MA: Blackwell.
- Greenberg, M. (2009). Moral concepts and motivation. *Philosophical Perspectives*, 23, 137–164.
- Greenberg, M. Incomplete understanding, deference, and the content of thought (manuscript).
- Hellie, B. (2004). Inexpressible concepts and the allure of the knowledge argument. In P. Ludlow, Y. Nagasawa, & S. Daniel (Eds.), *There’s something about Mary* (pp. 333–364). Cambridge: MIT Press.
- Horgan, T., & Tienson, J. (2002). The phenomenology of intentionality and the intentionality of phenomenology. *Pacific Philosophical Quarterly*, 84, 1–23.
- Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, 32, 127–136.
- Jackson, F. (1986). What Mary didn’t know. *The Journal of Philosophy*, 83, 291–295.
- Jackson, F. (1998). *From metaphysics to ethics: A defence of conceptual analysis*. Oxford: Oxford University Press.
- Levine, J., & Trogon, K. (2009). The modal status of materialism. *Philosophical Studies*, 145, 351–362.
- Lewis, D. (1988). What experience teaches. *Proceedings of the Russellian Society*, 13, 29–57.
- Lewis, D. (2002). Tharp’s third theorem. *Analysis*, 62, 95–97.
- Loar, B. (1990). Phenomenal states. *Philosophical Perspectives*, 4, 81–108.
- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83, 435–450.
- Nemirow, L. (1990). Physicalism and the cognitive role of acquaintance. In W. Lycan (Ed.), *Mind and cognition* (pp. 490–499). Cambridge: Blackwell.
- Nida-Rumelin, M. (1996). What Mary couldn’t know: Belief about phenomenal states. In T. Metzinger (Ed.), *Conscious experience* (pp. 219–242).
- Papineau, D. (1998). Mind the gap. *Philosophical Perspectives*, 12, 373–389.
- Perry, J. (2001). *Knowledge, possibility, and consciousness*. Cambridge: MIT Press.

- Putnam, H. (1975). The meaning of meaning. In Gunderson, K. (Ed.), *Language, mind and knowledge. Minnesota Studies in the Philosophy of Science* (Vol. 7, pp. 131–193). Minneapolis: University of Minnesota Press.
- Rabin, G. Physicalism and the idealization of a priori entailment (manuscript). <https://sites.google.com/site/gabrielrabin/site/home>.
- Schwarz, W. (2007). Conceptual metaphysics and modal metaphysics. In H. Bohnse, K. Dreimann, & W. Sven (Eds.), *Selected papers contributed to the sections of GAP.6, sixth international congress of the German society for analytical philosophy* (pp. 520–528). Paderborn: Mentis.
- Stoljar, D. (2005). Physicalism and phenomenal concepts. *Mind and Language*, 20, 469–494.
- Strawson, G. (1996). *Mental reality*. Cambridge: MIT Press.
- Tye, M. (1995). *Ten problems of consciousness*. Cambridge: MIT Press.
- Tye, M. (2000). *Color, consciousness, and content*. Cambridge: MIT Press.
- Tye, M. (2009). *Consciousness revisited: Materialism without phenomenal concepts*. Cambridge: MIT Press.