



Book Review

Our Knowledge of the Internal World, by Robert Stalnaker

Clas Weber
Uppsala University

Disputatio Vol. 3, No. 25

November 2008

DOI: 10.2478/disp-2008-0011

ISSN: 0873-626X

philosophical logic may also interest themselves with the logical appendices, one of which presents modal logic as a subsystem of the logic of counterfactuals. Last but not least, the work also includes an afterword that is both a severe reprimand to the analytic community for a certain sloppiness and an exhortation to all colleagues to apply more rigor and patience in addressing metaphysical issues. People familiar with Williamson's work will not be surprised by the careful and detailed (sometimes a bit technical) argumentation, which demands careful attention from the reader. As expected, this is a most relevant contribution to an increasingly popular topic by one of today's leading analytic philosophers.

Luis S. Robledo

Sociedad Argentina de Análisis Filosófico, Buenos Aires

LOGOS Research Group, Barcelona

lsrobledo@yahoo.com.ar

Our Knowledge of the Internal World, by Robert Stalnaker.
Oxford: Oxford University Press, Clarendon Press, 2008, 160 pp.

In spite of decades of extensive discussion, conscious experience remains puzzling, stubbornly resisting accommodation within a naturalistic picture of the world. In his recent monograph *Our Knowledge of the Internal World*, given as the John Locke Lectures at the University of Oxford in 2007, Robert Stalnaker tries to solve the puzzle by offering an account of conscious experience that locates it within a materialistic conception of reality — an account that does not assign the knowledge of our experiences a privileged and foundational epistemological status. The book paradigmatically exemplifies one of Stalnaker's most characteristic philosophical virtues: a rare combination of depth and breadth in perspective with formal rigor and technical precision. It is extraordinarily rich and dense, requiring a close reading, that is, however, more than rewarded.

The background theme for *Our Knowledge of the Internal World* is Stalnaker's externalism. This theme comprises a number of different motifs. One is an anti-foundationalist approach to epistemology: there is no privileged basis from which to build up our knowledge, we have to 'start in the middle.' A further element is a naturalistic reductionism about intentionality; one, moreover, according to which inten-

tional content is determined in part by our environment. Finally, Stalnaker's externalism is characterized by a form of contextualism: we ascribe intentional states, including experiential ones, to others in order to explain their rational capacities and dispositions and those ascriptions are performed from and determined by the attributor's external context. While these components form a compelling unit, it is not clear that the internalist is bound to an outright rejection of the whole package, and could not take advantage of one or the other of its attractive features. According to Stalnaker, however, only a thorough-going externalism offers the hope to solve the problems before us:

[...] Cartesian and traditional empiricist ideas that presuppose an internalist perspective continue to influence the way we think about [knowledge and the mind], and some of the puzzles about our knowledge of our own experience and thought may arise from equivocating between internal and external perspectives (p. 4).

From this externalist outlook, Stalnaker approaches the central part of the book, a discussion of Frank Jackson's knowledge argument, illustrated by the famous story of Mary. In chapter two, he reviews and rejects several reactions to Jackson's challenge to materialism. The Fregean strategy is found wanting, since it ultimately relies on a separation of an overarching realm of conceptual possibilities from an alleged subclass of it, the metaphysical possibilities. In contrast, Stalnaker claims, we should understand our conceptual capacities in terms of the divisions we are able to affect among a fundamental domain of metaphysical possibilities. Likewise, Laurence Nemirow's and David Lewis' ability hypothesis is dismissed, since the capacities Mary acquires seem to be *cognitive* in nature, ultimately endowing her with the disputed knowledge-that. John Perry's indexical theory is regarded as on the right path. In the end, however, his account in terms of *reflexive content* does not succeed either. For one, Perry locates what in Stalnaker's eyes belongs to content itself, within the *means* of representing content. Moreover, the crucial epistemological change might occur in Mary, while she is still ignorant of the corresponding reflexive information.

Even though he does not accept Perry's account as a whole, Stalnaker embraces the central analogy between phenomenal and self-locating knowledge. To get a clearer grasp on what self-locating

content is, he examines in chapter three Lewis' theory thereof. Finding it inadequate, he develops in detail his own conception.

Chapter four introduces Stalnaker's own response to Jackson's argument. The main idea is that phenomenal information is analogous to essentially self-locating information in a crucial respect: both can only be acquired from within a certain context. Therefore, Mary could not have possessed the relevant knowledge before leaving her room. His interpretation of the Mary case will in the end cast into doubt the assumption that we are intimately 'acquainted' with our experiences, our intuitions about the transparency of our thoughts, as well as the privileged epistemic role experiential knowledge is often supposed to play.

Taking once more Lewis as an exemplar, Stalnaker convincingly illustrates in chapter five the problems that result from rejecting the acquaintance picture of experience, while at the same time assigning phenomenal knowledge a special foundational epistemic role. We should, he argues, follow Lewis in giving up the claim that we know the intrinsic nature of our experiences, and consequently refrain from reserving for experiential knowledge a special evidential status: 'Our epistemic relation to our experience is like our epistemic relation to anything else in the world' (p. 93).

Demonstrating that such a position is compatible with a qualified version of the thesis that our own thoughts are transparent to us is the task of chapter six. The key is a consistent commitment to the externalism we set out with, especially its contextualist aspect. We ascribe beliefs and knowledge from an external point of view. And what 'counts as a correct description of the world according to the thinker may depend on the attributor's context' (p. 131). Furthermore, we are bound to ascribe knowledge and thought in a manner that maximizes the rationality — of which transparency of thought is an integral part — of the subjects we ascribe it to. As he emphasizes in the final chapter, Stalnaker takes this contextualist account of knowledge and thought to be compatible with a realist understanding of those notions.

Here, I want to concentrate on Stalnaker's discussion of the knowledge argument and his own indexical solution. In a rough and simplistic form the knowledge argument runs like this:

- (1) Mary knows all physical truths/facts.
- (2) However, she is ignorant of certain phenomenal truths/facts.
- (C) Therefore, there are truths/facts that are not physical truths/facts, and hence materialism is false.

His critique of the Fregean strategy demonstrates that Stalnaker does not think one could gain any leverage in distinguishing truths from facts. Moreover, he accepts that Mary lacks the relevant phenomenal knowledge, but refuses to endorse the conclusion that materialism is false. So, where does he object to the argument?

The assumed analogy between phenomenal and self-locating knowledge suggests the following interpretation. Perhaps, Stalnaker holds on to materialism by rejecting the following tacit premiss:

- (3) If there are facts that are not physical facts, then materialism is false.

To see why that premiss is problematic, consider the following indexical argument against materialism:

- (1) Mary knows all physical facts.
- (2) However, she is ignorant of certain indexical facts.
- (C) Therefore, there are facts that are not physical facts, and hence materialism is false.

Few would accept the indexical version of the knowledge argument. Consequently, we should distinguish between two kinds of facts: (i) heavyweight objective facts, and (ii) lightweight indexical facts. Only ignorance of facts of the former kind is inconsistent with materialism. The failure of our physical conception of reality to entail all subjective, indexical facts does not demonstrate that this conception leaves out something important about how 'the world is in itself.' Indexical facts, rather, concern our point of view within that objective reality. And so do phenomenal facts, Stalnaker suggests: 'facts about phenomenal experience should be understood as features of our perspective on the world' (p. 75). Therefore, we can react to the original argument in the same way we reacted to its indexical version.

Compelling as that interpretation may seem, I do not think it fully captures the crucial point of Stalnaker's account. As I understand Stalnaker, the phenomenal knowledge Mary acquires allows her to exclude a real heavyweight physical possibility. He rather locates the flaw of the knowledge argument in the assumption that Mary could possess *all* the relevant knowledge in a context-independent way. There are certain items of information that are *essentially contextual*. That is the important feature phenomenal knowledge shares with self-locating knowledge, and it is that feature that explains Mary's ignor-

ance. She just *could not* have had the knowledge beforehand. Accordingly, I take Stalnaker to question the intelligibility of premiss (1).

On that second interpretation, a lot hangs on the idea of essentially contextual information. What exactly is essentially contextual information and what is Stalnaker's reason for thinking there is such a thing?

In Stalnaker's framework, the meaning of an utterance is to be understood in terms of the effects it has on the context (understood as the set of possible worlds that correspond to the shared assumptions of the conversational partners). We might represent utterance-meaning as a function that maps prior contexts onto posterior contexts. In ordinary circumstances, such a function will be defined for a whole range of contexts: the meaning of 'pigs can fly' will subtract from arbitrary contexts the worlds in which pigs cannot fly. Now, as a limiting case, such a function from contexts to contexts might be defined for just one single context; then, 'the content cannot be detached from the context in which it is expressed or thought' (p. 81). Or in propositional terms: the relevant belief or utterance does not affect a division among the whole of logical space. The corresponding proposition is only defined for the set of alternative possibilities that are relevant in the local context. In such cases, the information is *essentially contextual*.

The positive case Stalnaker makes for the existence of essentially contextual information is rather thin. He illustrates it with a little example (here slightly simplified):

A and B are facing a building that covers all but the bow and stern of a large ship. B wrongly believes there to be two different ships. A, pointing first to the bow, then to the stern, utters: '*This* is the same ship as *that*.' A has thereby eliminated from B's belief-worlds a certain possibility.

Prima facie, it is difficult to semantically describe the effect of A's utterance, since we cannot define it in relation to two actually existing ships. In previous work, Stalnaker himself has provided a method to specify the content of A's utterance: diagonalization. The possibility A excludes is one in which her two pointings demonstrate different ships. Why does Stalnaker now think that this solution is insufficient? He seems to think that it does not give us the intuitively right information: A's utterance is not *about* the utterance context (the pointings) itself, but *about* the ship. It is not clear what the force of this intuition is. It follows from Stalnaker's own account of propositions

that two propositions that eliminate the same possibilities are identical, irrespective of the means we use to specify them, or our intuitions of *aboutness*. Admittedly, we can still distinguish the two propositions, since only one of them is defined outside of the present context. However, we wanted to have a reason for thinking they differ in that respect. And Stalnaker appeals to our intuitions about a difference in *information* or *aboutness* between the two, and not to the fact that we have no grip on what content the utterance conveys outside of the local context. He grants that we can *describe* that information from the outside, but maintains that this is not the information *expressed*. I find this contrast somewhat elusive and would have liked to hear more on why one should believe in essentially contextual information.

There is another problem: if phenomenal information is essentially bound to a local context, we cannot communicate it to subjects outside of that context. More importantly, we seem to lose the information ourselves, once we leave the context. Our phenomenal knowledge, one would think, is more stable than that.

Finally, Stalnaker's indexical examples of essentially contextual knowledge are all cases of demonstrative singular reference. However, the crucial experiential claim is one about *properties*: 'seeing red is like *this*', where '*this*' refers to an experience-*type*. It is not clear that the model easily translates from singular to general reference.

Let us move on to Stalnaker's modified Mary-scenario (again simplified):

In her room Mary is told that, depending on the flip of a coin, she will soon be in either of two situations: she will either be shown a red or a green object. Call the situations 'R' and 'G'. Mary knows everything physical about them. In fact, unbeknownst to her, she is shown a red object, i.e. she is in situation R. Still, Mary cannot exclude situation G as actual.

According to Stalnaker, the relevant phenomenal knowledge will ultimately allow Mary to exclude situation G, a heavyweight physical possibility. Importantly, just seeing the red object does not put her in a position to do so, Stalnaker argues, since she does not know that she is seeing something red. What changed about Mary's epistemic situation is that she can now *entertain* those possibilities. Stalnaker takes the described ignorance to be an important pre-theoretic datum, and it is crucial for his argumentation. It is Mary's ignorance of whether R or G is actual that casts into doubt the acquaintance picture of experience, the transparency of thought, and the principle of phenomenal

indistinguishability. In situation G, which she supposedly cannot exclude, Mary is having a different experience, undermining the principle that all her epistemic alternatives are phenomenally indistinguishable, as well as the claim that we are acquainted with the intrinsic nature of our experience. Under the assumption that there are such things as pure phenomenal concepts, she would employ a different pure phenomenal concept in G, threatening the internalist version of the transparency of thought thesis. Stalnaker suggests that we should go with the pre-theoretic datum, and not with contested theoretical assumptions. I want to question that datum.

As Stalnaker points out, R and G are ‘phenomenally as well as physically different’ (p. 90). Now, consider the following modification of Stalnaker’s story:

Mary is first shown something red, then something green. The first situation is called ‘1’, the second ‘2’. Then, she is once more shown a red object. Now we ask her: *Is your present situation like 1 or 2?*

I think it is intuitively obvious that Mary will be able to exclude being in a situation of type 2, i.e. she can eliminate situation G. Since she still does not know whether she has seen something red or something green, Stalnaker would predict the opposite. The flaw in Stalnaker’s reasoning seems to be to identify Mary’s ignorance of what it is to see a red object, on the one hand, with her ignorance of what it is like to have a red experience, on the other. In fact, Mary’s ignorance does not concern what red experiences are like. What she does not know is that she has seen a red object. Maybe Stalnaker is right that a materialist should in the end not make that distinction. That, however, is not a pre-theoretic datum.

I take my criticisms to undermine to a certain extent the positive case Stalnaker makes for his account. It might nonetheless be an attractive picture in its own right.

Clas Weber
Department of Philosophy
Uppsala University
Thunbergsvägen 3H,
752 68 Uppsala, Sweden
Clas.Weber@filosofi.uu.se