

## Infallibility, Acquaintance, and Phenomenal Concepts

Wolfgang BARZ<sup>†</sup>

### ABSTRACT

In recent literature, there is a strong tendency to endorse the following argument: There are particular judgments about one's current phenomenal experiences that are infallible; if there are particular judgments about one's current phenomenal experiences that are infallible, then the infallibility of those judgments is due to the relation of acquaintance; therefore, acquaintance explains why those particular judgments about one's current phenomenal experiences are infallible. The aim of this paper is to examine critically both the first and the second premise of this argument. It will emerge that there might be a small class of judgments about one's current phenomenal experiences that are infallible, namely judgments involving direct phenomenal concepts. However, as I will try to show, the infallibility of such judgments, if existent at all, is not due to the relation of acquaintance.

### *1. Introduction*

For a long time, it seemed as though the idea of infallibility was dead and buried for good. Nobody would ever argue again that it is impossible to have a false belief about one's current mental state. Times have changed, however. Since the turn of the millennium or so, the idea of infallibility is undergoing a renaissance. An increasing number of philosophers are willing to subscribe to some version or other of the infallibility thesis. This is especially conspicuous from the following selection of seminal contributions to the current debate about phenomenal concepts. Brie Gertler set the ball rolling when she claimed that certain introspective judgments about the

---

<sup>†</sup> University of Duisburg-Essen, Department of Humanities, Institute of Philosophy, 45117 Essen, Germany; Email: wbarz@zedat.fu-berlin.de.

phenomenal character of one's own current mental states<sup>1</sup> enjoy an epistemic asset she calls "Cartesian certainty" (Gertler 2001, 327). David Chalmers followed suit in declaring that what he calls "direct phenomenal beliefs" cannot be false (Chalmers 2003, 242). Then came Terry Horgan and Uriah Kriegel, who took a stand for "limited phenomenal infallibility" (Horgan & Kriegel 2007, 129; see also Horgan 2012, 409–412). The latest example is Katalin Balog. According to her "quotational account of direct phenomenal concepts," judgments of the form "I have R"—where R is a direct phenomenal concept—cannot fail to be true (Balog 2012, 27).<sup>2</sup>

---

<sup>1</sup> Let me clarify briefly the ontological framework that I will use in speaking about the mental. As Chalmers (2003, 222 f.) and Gertler (2012, 96 f.) explain, there are at least two different ways of thinking about the bearers of phenomenal properties. The one is to conceive of phenomenal properties as being instantiated by *subjects* (at certain times), the other one is to say that phenomenal properties are exemplified by the *experiences* the subject has (at certain times). These different ways of thinking suggest different views about experiences: one might think that experiences are nothing but instantiations of phenomenal properties, or one might think that experiences are more than that, namely, mental individuals that instantiate phenomenal properties. However, as both Chalmers and Gertler note, nothing depends on whether one thinks of phenomenal properties in the one or the other way (at least for the discussion at hand). Therefore, for the sake of being in accord with Balog (2012), whose account of phenomenal concepts I will criticize in the second part of this paper, I will treat phenomenal properties as though they are exemplified by experiences. For reasons of stylistic variation, I will sometimes also speak of the "phenomenal character" of an experience.

<sup>2</sup> The idea of infallibility also plays a prominent role in recent discussions about the nature of phenomenal consciousness. For it seems that the two most promising approaches to this issue—the higher-order and the self-representational approach—are equally plagued by problems that arise from the fact that representational states (be they of higher-order or of same-order type) might *misrepresent* their target states. This is an unwelcome consequence, since it is at odds with the common view that it is impossible to be conscious of a particular phenomenal property, while that property is not exemplified by any of one's current experiences. (See Neander (1998) for an exposition of this problem for higher-order approaches. Weisberg (2008) applies this problem to self-representational approaches.) Thus, proponents of both the higher-order and the self-representational approach face the task of somehow accommodating the common view. On behalf of the self-representational approach, for example, Chad Kidd (2011) recently argued for a thesis he calls "contingent infallibility." According to Kidd, "it is contingently impossible for self-representations ... to misrepresent their objects" (361). In order to be as clear as possible about the topic of my paper, let me say that Kidd's thesis of contingent infallibility should *not* be lumped together with the claim—advocated by Gertler, Chalmers, Horgan and Kriegel, and Balog—that certain introspective judgments about the phenomenal character of one's own current mental states are infallible (whether contingently or not). As Kidd himself aptly explains, his thesis of contingent infallibility does not refer to "reflective or introspective self-representations," but to "pre-reflective self-representations" (370). Therefore, even if both use the notion of infallibility, there is a clear difference between Kidd's thesis and the thesis advocated by Gertler et al. While the first concerns representational structures that are, according to the self-representational approach Kidd advocates, proper parts of the experiences whose phenomenal properties they represent; the latter concerns representational structures that, even according to Kidd's self-representational approach, do not form proper parts of the experiences they are about. Gertler et al. claim that certain *higher-order judgments* are infallible, but they are silent about the issue whether experiences have proper parts that infallibly represent the phenomenal character of the experience of which they are part. In order to forestall possible misunderstandings from the start, let me emphasize that my paper focuses exclusively on the infallibility thesis as maintained by Gertler et al.—and *not* as proposed by Kidd. (Yet I will come back to Kidd's thesis occasionally in order to clarify

While Gertler, Chalmers, Horgan and Kriegel, and Balog might disagree with respect to the exact formulation of the infallibility thesis, they all share the view that the key to the explanation of why certain judgments about one's current experiences are infallible lies in a certain feature of those experiences themselves. Horgan and Kriegel aptly express this view when they say that there "is something about phenomenal experiences that makes them amenable to safer, and ultimately under some conditions infallible, knowing" (Horgan & Kriegel 2007, 139).<sup>3</sup> This "something" is supposed to be the fact that *undergoing an experience* goes hand in hand with *being in a privileged epistemic position vis-à-vis that experience*. It seems that, at the very moment I have an experience, a special relation between my experience and me is established that somehow enables me to tell infallibly what the experience is like. Gertler (2012), Chalmers (2003) and Balog (2012)—borrowing from Russell—call this relation "acquaintance,"<sup>4</sup> whereas Horgan and Kriegel—more in line with Brentano—prefer to speak in terms of an "inner awareness" which is somehow "built into ... the experience it is an awareness of" (Horgan & Kriegel 2007, 132). All details aside then, Gertler et

---

matters.)

<sup>3</sup> Horgan and Kriegel advocate the self-representational approach to phenomenal consciousness according to which a conscious mental state, say the perceptual experience of something blue, is a structured whole that consists of at least two parts: first, a representation of a blue object—the "outer-representation," as it is called by Kidd—, second, a representation of the phenomenal character of the outer representation—the "self-representation," as Kidd calls it. So indeed, there is a point of contact between Kidd (2011), on the one hand, and Horgan & Kriegel (2007), on the other: the fact that a person, as Horgan & Kriegel (2007) claim, is "innerly aware" of the phenomenal character of its current experiences is due to the fact that each of her current experiences contains a component Kidd (2011) calls "self-representation." However, here the parallels end. In particular, Horgan & Kriegel's "limited phenomenal infallibility" claim neither implies, nor is implied by, Kidd's "contingent infallibility" claim. As I explained above, Kidd's claim concerns the question whether the "self-representation" might misrepresent the phenomenal character of the "outer-representation." This is an issue between two components *of one and the same mental state*, so to speak. In contrast, the claim of Horgan & Kriegel concerns the question whether the phenomenal character of a certain experience  $\varphi_1$  might be misrepresented *by another mental state*  $\varphi_2$  (which is not necessarily an experience) *that is directed at*  $\varphi_1$ . This, as it were, is an issue between *two different mental states*.

<sup>4</sup> Balog (2012, 16) provides a striking account of the idea: "Acquaintance is a unique epistemological relation that relates a person to her own phenomenally conscious states and processes *directly, incorrigibly, and in a way that seems to reveal their essence*. When one is aware of a phenomenal state in the process of having it, something essential about it is revealed, directly and incorrigibly—namely, *what it is like to have it*" (original emphasis).

al. seem to argue along the following lines: There are particular judgments about one's current phenomenal experiences that are infallible. If there are particular judgments about one's current phenomenal experiences that are infallible, then the infallibility of those judgments is due to the relation of acquaintance (or inner awareness). Therefore, acquaintance (or inner awareness) explains why those particular judgments about one's current phenomenal experiences are infallible.<sup>5</sup>

I frankly admit from the start that I feel dubious about this kind of argument. I see it as an unfortunate step backward in philosophy. It would be as though Neurath, Quine, and Sellars had never existed. However, I do not want to deny that there are judgments about one's own current experiences that are infallible. Though I doubt whether most of the introspective judgments that have been suggested hitherto as being infallible really are infallible, I concede that there might be a small class of introspective judgments, which are infallible indeed. Nor do I want to deny that there is a unique relation each of us bears to his or her own current experiences—acquaintance, if you like. However, what I want to deny is that acquaintance *explains* why the elements of that small class of introspective judgments are infallible. There might be a relation like acquaintance—I can agree with that. However, our infallibility on phenomenal matters, if there is any, is not *due to* this relation.

Here is a sketch of my argument in this paper: Proponents of the new wave of infallibilism typically restrict their claim to special judgments about the phenomenal character of one's own experiences—namely, judgments that involve so-called “direct phenomenal concepts.”<sup>6</sup> In order to gain a good grasp of those judgments in general,

---

<sup>5</sup> In this sense, Gertler (2011, 87) says: “Our acquaintance with mental states allows introspective self-knowledge to be especially secure, epistemically.” Chalmers (2003, 248) puts the same idea as follows: “So acquaintance is a relation that makes this sort of lucid understanding [a certain sort of lucid understanding of phenomenal properties that is embodied in pure phenomenal concepts] possible.”

<sup>6</sup> This is Chalmers' (2003) term. Balog (2012) likewise uses this term. Gertler (2012), in contrast, prefers to speak of “introspective demonstratives.” Horgan and Kriegel, in turn, use the notion of a “bracketing

and of direct phenomenal concepts in particular, I would like to reconstruct the logical route from the unrestricted to the restricted thesis as it is advocated by contemporary adherents of infallibility.<sup>7</sup> Thus, in Section 2, I will start out from the simple idea that all judgments about one's own current mental states have to be true and, on this basis, develop more and more restricted infallibility claims, each of which will turn out to be unconvincing.

In Section 3, I introduce the claim that I am primarily interested in, namely, the idea that particular introspective judgments that involve tokenings of so-called “direct phenomenal concepts” cannot be false (“Infallibility<sub>6</sub>”). I reconstruct this claim as a response to David Armstrong's seminal “distinct existences” argument, according to which an experience and an introspective judgment about this very experience are distinct entities and therefore can exist independently from each other. Introspective judgments that involve direct phenomenal concepts provide counterexamples to Armstrong's claim. Since direct phenomenal concepts are built from the experiences to which they refer, experiences somehow *form parts* of introspective judgments that involve direct phenomenal concepts—and, thus, are not strictly distinct from those judgments. In order to put the idea of direct phenomenal concepts across a little more specifically, I draw on Katalin Balog's (2012) “quotational account,” according to which a direct phenomenal concept contains an occurrence of the experience to which it refers in roughly the same way as a quotation contains a token of the word or phrase to which it refers. I will try to show that, even if Balog's account of direct phenomenal concepts is correct, “Infallibility<sub>6</sub>” is false.

---

mode of presentation” (Horgan & Kriegel 2007, 128).

<sup>7</sup> In this respect, others have already said many important things. See, for example, Horgan & Kriegel (2007), who develop their restricted infallibility thesis by criticizing less restricted claims. See also Gertler (2011, 70-80) who gives an excellent account of the “limits to the domain of privileged access.” However, I hope that my account will shed light on some aspects of the discussion that have been neglected thus far.

However, I will not leave it at that. In Section 4, I present a version of the infallibility claim that is immune to any counterexample (“Infallibility<sub>7</sub>”). What initially appears to give proponents of infallibility an edge turns out to be rather to their detriment upon further examination. I argue that “Infallibility<sub>7</sub>” does not hold because we are acquainted with phenomenal experiences, but rather that it holds due to a peculiar interplay between the existence condition and the truth condition of those judgments that “Infallibility<sub>7</sub>” is about—a phenomenon I call *self-verification*. Thus, I will conclude that the principal tenet of contemporary adherents of infallibility—that acquaintance (or inner awareness) *explains* why introspective judgments of a certain kind are infallible—is simply false.

## 2. *The logical route from Infallibility<sub>0</sub> to Infallibility<sub>5</sub>*

As a starting point, let us consider the infallibility<sup>8</sup> thesis in its unrestricted form:

Infallibility<sub>0</sub>: Necessarily, if *S* judges “I am in  $\varphi$  now” at *t*—where the concept of  $\varphi$  refers to a mental state type—, *S* is in  $\varphi$  at *t*.<sup>9</sup>

---

<sup>8</sup> Instead of infallibility, some authors speak of “incorrigibility” or even “indubitability” here. See, for instance, Armstrong (1968) and Jackson (1973).

<sup>9</sup> Two remarks. First, I prefer using “to judge” instead of “to believe” since beliefs are normally understood to be *dispositional* mental states. The second-order mental state whereby one becomes aware of one’s own first-order mental state, however, is of an *episodic* nature. In my view, this is best captured by the verb “to judge.” Second, I use direct instead of indirect speech construction in specifying the content of a judgment because indirect speech forces us to shift indexicals and, thus, might obscure the original content of a given judgment. For example, suppose that, on 6 October 2015 at noon, President Obama judges something that, if he wished to do so, he himself would have expressed by uttering the sentence “I am in pain now.” Now, suppose that I report this mental episode on 7 October 2015 in the evening. Using indirect speech, I say: “Yesterday at noon, President Obama judged that he was in pain then.” This report, however, leaves open the exact content of Obama’s original judgment, since the report is true even if, on 6 October 2015 at noon, Obama had judged something that he would have expressed then by uttering the sentence “On 6 October 2015 at noon, the 44th President of the United States is in pain.” In contrast, a report that uses direct speech will not be subject to this kind of indeterminacy: “Yesterday at noon, President Obama judged ‘I am in pain now’” could not be true if, on 6 October 2015

*Infallibility*<sub>0</sub> faces an objection that might not be obvious. Suppose that the slight headache I feel occasionally is identical to a certain state of my brain—call it neural state *N*. Suppose furthermore that, in order to detect my own brain states, I make use of an MRI scanner that, unbeknownst to me, is broken, and therefore falsely indicates that I am in neural state *N*. Since I believe that the MRI scanner is functioning properly, I make the judgment “I am in *N* now.” Here we have a case in which *S* judges “I am in  $\varphi$  now” at *t*—where the concept of  $\varphi$  refers to a mental state type, namely, the slight headache *S* feels from time to time—but *S* is *not* in  $\varphi$  at *t*.

One may object that, since a slight headache will never ever be identical to a brain state, the whole scenario is flawed. This objection, however, is not recommended, for thereby one would support one thesis, *Infallibility*<sub>0</sub>, by another one, anti-materialism, which is as controversial as the first. I think, therefore, that the best move available to the proponent of infallibilism in the face of the MRI-objection is to withdraw *Infallibility*<sub>0</sub> and fall back on a slightly modified thesis:

*Infallibility*<sub>1</sub>: Necessarily, if *S* judges “I am in  $\varphi$  now” at *t*—where “ $\varphi$ ” is a folk-psychological concept referring to a mental state type—, *S* is in  $\varphi$  at *t*.

I am not able to give a cut and dried definition of “folk-psychological concepts,” nor do I intend to claim that the extension of this term is clear-cut. I think, however, that a few examples are enough to make sufficiently clear what the idea is. There are admissible

---

at noon, Obama had judged something that he would have expressed then by uttering “On 6 October 2015 at noon, the 44th President of the United States is in pain.” – To be sure, this is not to say that the use of direct speech constructions in reports of judgments is utterly unproblematic. First and foremost, there may be doubts about the idea that we judge *sentences*—which seems to be presupposed by the use of direct speech constructions. It is not sentences, but *propositions* that we judge, one might say. However, I do not think that the use of direct speech constructions in reports of judgments presupposes that we judge sentences instead of propositions—it only presupposes that contents of judgments can be specified with the help of sentences. Thanks to an anonymous referee for pressing me on this point.

and inadmissible instances of “I am in  $\varphi$  now.” Examples of admissible instances are: “I am in pain now,” “I currently believe that tomorrow is Tuesday,” “I hear raindrops falling on the windowpane now,” etc. In contrast, the following instances are inadmissible: “My C-fibers are firing now,” “There is a token of ‘Tomorrow is Tuesday’ in my belief-box now,” “There is neural activity in the Brodmann areas 41 and 42 of my brain now,” etc.

However, even if we grant the notion of a folk-psychological concept, *Infallibility*<sub>1</sub> is subject to counterexamples. Consider cases in which a subject bases her judgment on the testimony of another person, on the observation of her own behavior or on neurophysiological evidence. Since such judgments are—as it is generally taken for granted—not immune to error, the infallibility thesis should be strengthened by restricting its scope to judgments which are based *on introspective grounds alone*. Let us call those judgments “introspective judgments” and the activity whereby one makes an introspective judgment “to judge introspectively.” Then we arrive at the following thesis:

Infallibility<sub>2</sub>: Necessarily, if *S* judges introspectively “I am in  $\varphi$  now” at *t*—where “ $\varphi$ ” is a folk-psychological concept referring to a mental state type—, *S* is in  $\varphi$  at *t*.

When is a judgment introspective and when is it not? Certainly, we should *not* say: the judgment “I am in  $\varphi$  now” is introspective if, and only if, it is infallible—since *Infallibility*<sub>2</sub> would be trivially true then, and not worth discussing. Subscribing to one of the many substantial accounts of introspection currently available is not an option either. Some of them regard the thesis that we have privileged access to (some of) our own mental states as an *explanandum*, whereas the others are built on the assumption



that we stand vis-à-vis our own mind just as we stand vis-à-vis objects and persons in our environment.<sup>10</sup> So, substantial accounts of introspection typically either *imply* or *rule out* that our epistemic position vis-à-vis our own mental states is better than our epistemic position vis-à-vis other things in the empirical world. Thus, subscribing to one of those substantial accounts of introspection would mean to take a biased viewpoint towards the issue of infallibility. Therefore, in order to remain on neutral ground, I prefer to be silent on the exact nature of introspection. Instead, I propose the following working criterion: A judgment is introspective if (1) it can be expressed by sentences of the form “I am in  $\varphi$  now”—where “ $\varphi$ ” is a folk-psychological concept referring to a mental state type—and (2) is not based on any conscious inference. In order to exclude objections to the effect that this criterion cannot rule out cases in which, for example, nefarious neuroscientists manipulate the brain of some unfortunate person in such a way that will induce her to make non-inferential judgments of the form “I am in  $\varphi$  now” when they push a button, one might add a third condition, namely, that a judgment is not introspective unless (3) it rests on something like attention to one’s current mental states.

Even *Infallibility*<sub>2</sub>, however, seems to be open to counterexamples. Consider cases of deep-rooted prejudices. Here is an example from Eric Schwitzgebel (2012):

Many men in academia sincerely profess that men and women are equally intelligent. Ralph, a philosophy professor let us suppose, is one such man. He is prepared to argue coherently, authentically, and vehemently for equality of intelligence and has argued the point repeatedly in the past. And yet Ralph is systematically sexist in his spontaneous reactions, judgments, and unguarded behavior. When he gazes out on the class the first day

---

<sup>10</sup> The latter is true of the accounts of Ryle (1949) and, to a lesser extent, Armstrong (1968) and Lycan (1996). The former applies to so-called “constitutive” accounts like Shoemaker’s (1994), “neo-expressivist” accounts like Bar-On’s (2004) as well as to acquaintance accounts like Gertler’s (2001).

of each term, he cannot help but think that some students look brighter than others—and to him, the women rarely look bright. When a woman makes an insightful comment or submits an excellent essay, he feels more surprised than he would were a man to do so, even though his female students make insightful comments and submit excellent essays at the same rate as his male students. When Ralph is on the hiring committee for a new office manager, it will not seem to him that the women are the most intellectually capable, even if they are; or if he does become convinced of the intelligence of a female applicant, it will have taken more evidence than if the applicant had been male. And so on. (Schwitzgebel 2012, 191 f.)

As Schwitzgebel’s description suggests, Ralph does *not* believe that men and women are equally intelligent.<sup>11</sup> Yet it does not seem inconceivable that Ralph introspectively judges “I believe that men and women are equally intelligent.” For he might be “disadvantaged by a desire not to see himself as sexist and by the more general desire to see himself as someone whose actions reflect his espoused principles” (Schwitzgebel 2012, 192).

Therefore, the proponent of infallibilism would be in a better position if she restricted the scope of the infallibility thesis even further, namely, to introspective judgments about *phenomenally conscious mental episodes*:

Infallibility<sub>3</sub>: Necessarily, if *S* introspectively judges “I am in  $\varphi$  now” at *t*—where “ $\varphi$ ” is a folk-psychological concept referring to a phenomenally conscious mental episode—, *S* is in  $\varphi$  at *t*.

---

<sup>11</sup> One might object that Ralph actually believes that men and women are equally intelligent, but just tends not to act on that belief. However, I follow Schwitzgebel in maintaining that “this seems to me an overly linguistic and intellectualist view of belief. Our beliefs manifest not just in what we say, but in what we do—they animate our limbs, not just our mouths—and they are also manifested in our spontaneous emotional reactions and our implicit assumptions” (Schwitzgebel 2012, 192). However, Schwitzgebel would not go as far as me in claiming that Ralph does *not* believe in the intellectual equality of the sexes. According to Schwitzgebel, Ralph is in an “in-between state” regarding the question whether men and women are equally intelligent.

*Infallibility*<sub>3</sub> faces difficulties as well. Many folk-psychological concepts referring to phenomenally conscious mental episodes are so structured that their application depends on the presence of certain external conditions about which the subject can be mistaken. This is apparent in connection with the concepts underlying the verbs “to see,” “to hear,” “to smell,” “to taste” and “to touch.” Nevertheless, it is also true of the concept of pain. It is generally accepted that the folk-psychological concept of pain is partly defined by a certain causal role: if a subject is in pain, then she is in a state that is, *inter alia*, typically caused by tissue damage and typically provokes the desire to get rid of this unpleasant sensation. One could imagine, then, a situation in which neuroscientists manipulate parts of my brain and thereby evoke a certain mental episode that is phenomenally indistinguishable from pain. Suppose furthermore that the neuroscientists are surprised when I report a feeling of pain, since they are manipulating a previously unknown part of the brain, say area *XY*, which is normally not involved in the production of pain. Let us even suppose that it is a law of nature that tissue damage *never* causes activity in area *XY*, and activity in area *XY* *never* provokes the desire to get rid of something. Therefore, the mental episode I undergo in this situation is not pain. (Instead, the neuroscientists speak of “prain” for fun.)

From the perspective of the proponent of infallibilism, the appropriate reaction to this challenge is to point out that pain and prain have something in common. They may be different with regard to causal role, but they undoubtedly are identical with regard to *phenomenal character*. Let us call the phenomenal character shared by pains and prains “ $Q_{\text{Pain}}$ .” It might be, then, that I can be mistaken about whether I am in pain now—but can I ever be mistaken about whether I am experiencing  $Q_{\text{Pain}}$  now? If I cannot be mistaken about this, the following thesis suggests itself:

Infallibility<sub>4</sub>: Necessarily, if *S* introspectively judges “I am experiencing *Q* now” at *t*—where “*Q*” is a phenomenal concept—, *S* is experiencing *Q* at *t*.

What is “phenomenal concept” supposed to mean here? The least that could be said is that phenomenal concepts are those concepts that are used to attribute phenomenal properties. As it is often remarked, there are no expressions in natural language that could be used exclusively for that purpose. Thus, the usual way to attribute a phenomenal property in natural language is to borrow from the stock of terms with which we normally refer to perceptible properties. In this sense, “red” has a double use: as a *perceptual* concept, it refers to the redness of, say, the Chinese flag. As a *phenomenal* concept, however, it refers to what it is like to have the visual experience of a Chinese flag—or, to put it in Byrne’s terms, it refers to that property that types my experience of a Chinese flag according to what it is like to undergo it.<sup>12</sup> Usually, the context clarifies to which property a speaker refers by using “red.”

However, even *Infallibility*<sub>4</sub> is not tenable. The reason is that it is susceptible to the problem of the speckled hen: When one sees a hen with 48 speckles on the side facing towards oneself in optimal conditions and, consequently, has a visual experience that exemplifies the phenomenal property of “48-speckledness,” one might be easily mistaken about the phenomenal character of one’s visual experience. For example, one may be incapable of introspectively discriminating between phenomenal 48-speckledness (“*Q*<sub>48</sub>” in short) and phenomenal 47-speckledness (“*Q*<sub>47</sub>”), and therefore makes the incorrect judgment “I am experiencing *Q*<sub>47</sub> now.”<sup>13</sup>

---

<sup>12</sup> Cf. Byrne (2002, 7): “the phenomenal character of an experience *e* is a *property*, specifically a *property of e*: that property that types *e* according to what it’s like to undergo *e*.” (I owe this quotation to Fish (2010, 17).)

<sup>13</sup> Cf. Chisholm (1942). For a recent formulation of the problem, see Gertler (2011, 103-104).

One might object that the addition of a single speckle makes no difference to the overall phenomenology involved in seeing the hen. So the judgment “I am experiencing  $Q_{47}$  now” is not at all incorrect, since the phenomenal property of 47-speckledness is *just the same as* the phenomenal property of 48-speckledness. However, this is unconvincing. If one accepts that  $Q_{47}$  is just the same as  $Q_{48}$ , then one seems to be committed to the view that  $Q_{48}$  is just the same as  $Q_{49}$ , for symmetrical reasons. If one iterates this move often enough, one obtains the thesis that  $Q_{999}$  is just the same as  $Q_{1000}$ . Now, recall that the relation of identity is transitive: if  $a = b$  and  $b = c$  then  $a = c$ . This implies that  $Q_{47}$  is just the same as  $Q_{1000}$ . However, this is absurd—for the overall phenomenology involved in seeing a hen with 47 speckles is very different from the overall phenomenology involved in seeing a hen with 1000 speckles.

Thus, if one still wants to hold on to the idea of infallibility, one has to restrict its scope once again:

Infallibility<sub>5</sub>: Necessarily, if  $S$  introspectively judges “I am experiencing  $Q$  now” at  $t$ —where “ $Q$ ” is a simple phenomenal concept—,  $S$  is experiencing  $Q$  at  $t$ .

When are phenomenal concepts “simple”? Phenomenal concepts are simple just if they refer to simple phenomenal properties. Examples of simple phenomenal properties are the usual suspects: phenomenal redness, phenomenal hardness, phenomenal warmth, phenomenal bitterness, pain (in its phenomenal core sense, i.e. detached from its causal role) etc.—but *not* properties like phenomenal 48-speckledness.

Admittedly, *Infallibility*<sub>5</sub> has some plausibility. It seems hard to imagine how a subject who makes a judgment to the effect that she herself currently experiences, say, the phenomenal property shared by pains and prains, could ever be mistaken. How,

then, might *Infallibility*<sub>5</sub> be defeated? As far as I can see, David Armstrong put the strongest argument against *Infallibility*<sub>5</sub> forward:

The acquiring of introspective knowledge must consist of the making of (sincere) reports of current mental occurrences, or else a nonverbal apprehension of these occurrences. In both cases the apprehension of the occurrence will have to be *distinct* from the occurrence that is apprehended. But if this is granted, then we can apply Hume's argument about 'distinct existences.' Wherever we have two distinct things, Hume points out, there we can always conceive of the one existing in the absence of the other. It follows that it is logically possible to have a sincere report of a current inner experience, or a nonverbal apprehension of that experience, without the experience existing. (Armstrong 1963, 422)<sup>14</sup>

Though I think that this is a good argument against familiar versions of the infallibility thesis such as *Infallibility*<sub>5</sub>, I beg to differ with Armstrong's wording here. The term "apprehension" is often used to refer to the relation that a subject bears to an experience just by undergoing it.<sup>15</sup> Nevertheless, this is not the meaning of "apprehension" as it is used by Armstrong. As Armstrong uses it, "apprehension" refers to that mental state whereby the subject *realizes* (or seems to realize), that she has an experience of a certain sort. This distinction between two meanings of "apprehension" is analogous to Ernest Sosa's distinction between what he calls *experiencing-awareness*, and *noticing-awareness*. The first is, as Sosa puts it, an awareness, "whereby one is 'aware' directly of an experience of a certain specific sort simply in virtue of undergoing it" (Sosa 2003, 277), whereas the second is an "intellectual awareness, whereby one occurrently believes or judges the thing noticed to be present, as characterized a certain way" (Sosa

---

<sup>14</sup> Armstrong attributes this argument to Smart (1962, 69). A version of this argument also appears in Armstrong (1968, 106 f.) – The Armstrong/Smart argument, to be sure, is not only aimed at defeating *Infallibility*<sub>5</sub>, but is intended to be used against all infallibility claims whatsoever. The fact that I present it as an argument against *Infallibility*<sub>5</sub> is just an effect of my paper's mode of presentation.

<sup>15</sup> See, for example, Lewis (1946, 188).

2003, 276). Therefore, Armstrong, instead of using “introspective apprehension,” might better have used “introspective judgment.” Armstrong’s argument, then, might be restated as follows:

- I. If two existent things are distinct, then it is possible that the one exists and the other does not.
- II. Even if *S* correctly judges “I am experiencing *Q* now” at *t*, this introspective judgment is still distinct from the phenomenally conscious experience it is about.
- III. Thus, it is possible that *S* introspectively judges “I am experiencing *Q* now” at *t*, though *S* does not experience *Q* at *t*.

In contrast to the attacks against versions of the infallibility claim considered above, Armstrong’s argument does not rely on more or less everyday-life counterexamples. Rather, it gives a highly theoretical reason to reject infallibility. Moreover, Armstrong’s argument can be seen as one of the main targets of contemporary proponents of infallibility like Gertler, Chalmers, Horgan and Kriegel, and Balog. The idea that experiences somehow “go into” (Horgan & Kriegel 2007, 138) some judgments in virtue of direct phenomenal concepts—an idea common to all those authors—poses a serious threat to Armstrong’s second premise. Thus, I will not proceed by directly formulating a more restricted infallibility claim that escapes Armstrong’s argument. Rather, I would first like to provide some *prima facie* justification of Armstrong’s argument to prepare for the idea of direct phenomenal concepts. Only then I will explain this idea in more detail and address the claim that I am primarily interested in: the claim that some introspective judgments that involve tokenings of direct phenomenal concepts are guaranteed to be true.

In order to get a better understanding of Armstrong's argument, let us take a closer look at its premises. First, notice that the idea of distinctness, as it is used in the first premise, should not be rendered as the negation of identity. The concept of identity, as it is usually understood, is governed by Leibniz's Law, which says that if there is something that has a property that something else lacks, then these objects are not identical. The crucial point is that the mere fact that two things are not identical does not support the claim that the one might exist without the other. Consider, for example, a wooden cross on top of a rocky hill. It is obvious that the cross as a whole is not identical to, say, its horizontal bar. For the mass of the horizontal bar is different from the mass of the cross as a whole. Nevertheless, it is also obvious that the cross could not exist without its horizontal bar: if we removed the horizontal bar, there would no longer be a cross on top of the hill, merely a pole. Therefore, in order to make the first premise plausible, the rendering of "x is distinct from y" has to be stronger than merely "x is not identical to y." Instead, it amounts to something like "x and y are neither identical, nor do they overlap or form part of each other."

In order to see the force of the second premise of Armstrong's argument, it might be helpful to recall a traditional difference that is often expressed by dichotomous conceptual pairs like "sensation and thought," "experience and judgment" or "perception and knowledge." These notions are commonly used to stand for two claims—one negative and the other one positive. The negative claim is that even if an object that exemplifies the perceptible property of being *F* presents itself in all its abundance, an intuition<sup>16</sup> of this object is not sufficient for judging—let alone for knowing—that it is *F*. The positive claim is that in order to judge that an object is *F* one has to apply the concept of *being F* to it. A first step of justifying the second premise,

---

<sup>16</sup> Here, "intuition" is to be understood in the sense of the German term "Anschauung."



then, might be to say that what is true of judgments about the external world has to be true of judgments about one's current mental states as well. One cannot judge, let alone know, that one is experiencing a certain phenomenal property unless one applies appropriate concepts to one's current conscious experience.<sup>17</sup>

Now, why is it that the fact that judgment presupposes the application of concepts supports the idea that introspective judgments about one's current conscious experiences are *distinct* from those experiences? There are two claims that are crucial here: The first is that, in a sense, judgments are exclusively composed of concepts; the second is that concepts are usually distinct from the things to which they refer. Let me briefly explain these claims in turn as they presuppose a certain view of how judgments are physically realized.

Following Frege, judgments may be seen as acts of recognition of the truth of a proposition.<sup>18</sup> This conception, however, leaves in the dark how judgments are realized by material creatures like us—for propositions are abstract entities that, due to their abstractness, do not causally interact with material stuff. Thus, it has become standard practice—at least in naturalistic circles—to conceive of judgments as tokenings of mental representations. According to the standard picture, mental representations have an inner structure that resembles the structure of sentences. For example, the sentence “My father is nice” might be parsed into the singular term “my father” and the predicate “is nice.” The mental representation that is tokened when one judges that one's father is nice likewise consists of parts that correspond to “my father” and “is nice.” These parts might be regarded as expressions of a language of thought—mental words, so to speak.

---

<sup>17</sup> Cf. Armstrong (1963, 422): “Introspective apprehension [...], like all other apprehension, is an apprehension that the thing apprehended is *of a certain sort*. The apprehension involves *classifying* the experience, in however rudimentary a way; that is, it involves the application of *concepts*.”

<sup>18</sup> Cf. Frege (1956, 294).

Mental words, in turn, may be called “concepts.”<sup>19</sup> Thus, the idea that judgments are exclusively composed out of concepts makes good sense—if one presupposes a certain view about the physical realization of judgments. I am well aware that this view is contestable. Yet I will accept it in what follows, since it is accepted by most of the participants in the debate that I address in this paper. Consider my commitment to the view in question as an assumption for the sake of argument—not as a central tenet of mine.

The claim that concepts are distinct from the things they refer to can be explained by using the analogy of natural language expressions. Consider the general term “tree.” This term is of a completely different nature than trees: it does not consist of bark, wood, and foliage, but of letters or sounds. Roughly the same can be said of the concept *being a tree*. Given the view of concepts that was put forward in the preceding paragraph, the concept *being a tree* likewise does not consist of arboreal stuff, but, presumably, of low-level syntactic units of the language of thought. Therefore, generalizing from this case, it might be maintained that there is no intersection between concepts, on the one hand, and the things they refer to, on the other. No concept is identical to the things to which it refers, and every concept is such that it does not overlap with the things to which it refers in the sense that one of those things forms part of it.<sup>20</sup> This, together with the claim that judgments are exclusively composed of concepts, implies that introspective judgments about the phenomenal character of one’s current conscious experiences are *distinct* from those experiences. Introspective

---

<sup>19</sup> Note that this use of “concept” deviates from the more common use according to which singular terms are not considered concepts. “Concept” in the sense stipulated here just means “word of the language of thought.” Consequently, even mental words which serve as singular terms qualify as concepts.

<sup>20</sup> A possible exception is the concept *being an expression of the language of thought*. Given that concepts are expressions of a language of thought, the concept *being an expression of the language of thought* is identical to one of the things to which it refers.

judgments may involve *conceptions* of those experiences, but not those experiences themselves—or so it seems.

### 3. *Direct Phenomenal Concepts and Infallibility*<sup>6</sup>

Having argued—as best I can—for the premises of Armstrong’s argument, I will now move on to assess those premises from the standpoint of the friends of infallibility. Some proponents of infallibility have attacked the first premise by pointing to distinct things that nevertheless seem to be necessarily tied to each other. In this sense, Frank Jackson writes that a “husband is numerically distinct from his wife, but ‘I am a husband’ entails ‘I have a wife’; an object’s colour is distinct from its extension, but ‘*A* is coloured’ entails ‘*A* is extended’; and so on” (Jackson 1973, 58; see also Raff 1966, 73). However, whatever the merits of this strategy, it is not the strategy of contemporary proponents of infallibility like Gertler, Chalmers, Horgan and Kriegel, and Balog. In contrast to Jackson, they could be understood as attacking, not the first, but the second premise of Armstrong’s argument. They claim that there is a subset of introspective judgments that *literally contain* the conscious experiences they are about. Brie Gertler nicely touches on this idea by speaking, concerning those judgments, of an “intersection between phenomenal reality and the epistemic” (Gertler 2012, 94).<sup>21</sup> If there is such an

---

<sup>21</sup> In the same spirit, Horgan & Kriegel (2007, 138) say that “there is something about phenomenal experiences that makes them ‘go into’ (some) beliefs about them.” – For the sake of clarity, allow me to compare the idea that certain introspective judgments literally contain phenomenal experiences with the idea promoted by Kidd (2011) and other proponents of self-representational approaches to phenomenal consciousness. Consider a conscious experience *M*. According to the self-representational approach, this very experience, *M*, has a part—which Kidd calls “self-representation”—that represents the phenomenal character of *M*. According to Gertler et al., *M* forms part of certain judgments about the phenomenal character of *M*—that may be likewise called “self-representations.” These are *different* (but, to be sure, not incompatible) theses: the first says that the “self-representation” (in Kidd’s sense) forms part of *M*, whereas the second says that *M* forms part of the “self-representation” (rendered as “judgment about the phenomenal character of *M*”).

intersection, i.e., if the experience that instantiates a phenomenal property forms a proper part of the judgment that one is currently experiencing this very phenomenal property, then of course, judgment and experience are not distinct from each other—at least not distinct in a sense that allows the judgment to exist without the experience it is about.

What are those introspective judgments, where “phenomenal reality and the epistemic intersect,” as Gertler puts it, supposed to be? Though there might be subtle differences from author to author, the core idea is that in making those introspective judgments the subject applies what David Chalmers calls a “direct phenomenal concept” to one of her current conscious experiences. Now, what is a direct phenomenal concept supposed to be? Chalmers says that the referent of a direct phenomenal concept “is somehow present inside the concept’s sense” (Chalmers 2003, 233). This is somewhat unsatisfying, however, since the sense of a concept is an abstract entity that lacks spatial structure. Thus, nothing can literally be *inside* a concept’s sense.

Katalin Balog has provided a more robust account of what a direct phenomenal concept is. She regards direct phenomenal concepts as analogous to linguistic expressions formed with the help of quotation marks.<sup>22</sup> In this sense, Balog claims that there is a concept-forming mechanism, called *mental quotation*, “that operates on an experience and turns it into a phenomenal concept that refers to either the token experience, or to a type of phenomenal experience that the token exemplifies” (Balog 2012, 34). Balog even uses a characteristic symbol for *mental quotes*, namely “\*,” in order to emphasize the analogy between linguistic and mental quotation. The only difference, she says, between linguistic and mental quotation “is that, unlike linguistic

---

<sup>22</sup> Cf. Balog (2012). See also Balog (2009). Balog is not the founder of the quotational account. It might be said that the quotational account originated with Lehrer (1997, 167–170). Balog’s source, however, is Papineau (2002, 116–121), who later distanced himself from the quotational account (cf. Papineau (2007, 120–123)).

quotation, what is between mental quotes ... is not a mental word but ... an experience” (Balog 2012, 35). So, the direct phenomenal concept of, say, phenomenal redness might be rendered as “\*◇\*”—where the diamond stands for some current conscious experience.<sup>23</sup> Thus, in light of Balog’s account, there is a clear sense in which a conscious experience is *inside* a concept: just as the quoting expression “‘tree’” literally contains a token of the word “tree,” the direct phenomenal concept “\*◇\*” literally contains a token of a “reddish” experience.<sup>24</sup>

Though Balog’s account is very clear with respect to the notion of an experience being *inside* a concept, one may have reservations about the idea of mental quotation. First and foremost, it seems to be unclear how a linguistic device such as quotation marks can be transferred to the case of mental representation. However, quotation is not essentially tied to the use of quotation marks. There are other devices that may be used instead. It may be, then, that Balog’s talk of *mental quotes* is simply a metaphor for some mental operation that takes a token experience and incorporates it into some complex whole—not necessarily comprising mental analoga to linguistic quotation marks—that serves as a concept of this very token experience (or of its phenomenal kind).<sup>25</sup> Since nothing in my argument depends on whether mental analoga to linguistic quotation marks exist, I will leave it at that. Note, moreover, that my aim here is not to *vindicate* Balog’s quotational account of direct phenomenal concepts. I am merely

---

<sup>23</sup> In typical cases, the experience between mental quotes exemplifies the phenomenal property to which the direct phenomenal concept refers. However, there may be exceptional cases where the experience between mental quotes *does not* exemplify the phenomenal property to which the direct phenomenal concept refers. See the discussion below for examples.

<sup>24</sup> “Reddish experience” is Balog’s term of art for an experience that instantiates the phenomenal property of redness; cf. Balog (2012, 27).

<sup>25</sup> An anonymous referee has even supposed that mental quotation might be nothing more than selective attention to the phenomenal character of experience as such plus the deployment of demonstratives in “inner ostension.” However, if mental quotation were nothing more than that, then we would lose the idea of introspective judgments that literally contain the conscious experiences they are about. Thus, I do not think that identifying mental quotation with selective attention plus deployment of demonstratives does justice to the quotational account.

interested in the consequences that would arise for the idea of infallibility *if the quotational account were true*.

In light of Balog's view of direct phenomenal concepts, Gertler's notion of "intersections between phenomenal reality and the epistemic" proves to be essentially correct. It seems that there are introspective judgments that *literally contain* the conscious experiences they are about—namely introspective judgments that involve direct phenomenal concepts. Thus, there are introspective judgments that are *not* distinct from the experiences they are about. Consequently, the second premise of Armstrong's argument is simply false.

However, the fact that the second premise of Armstrong's argument is false does not imply that *Infallibility*<sub>5</sub> is correct as it stands. *Infallibility*<sub>5</sub> seems to be too strong in that it ranges over *all* introspective judgments of the form "I am experiencing *Q* now"—where "*Q*" is some simple phenomenal concept. The reason is that there are introspective judgments of the form "I am experiencing *Q* now" that are distinct from the conscious experiences they are about, namely judgments that involve simple phenomenal concepts *that are not direct*. Therefore, an opponent to *Infallibility*<sub>5</sub> might rescue Armstrong's argument by repairing its second premise along the following lines:

Even if *S* correctly judges "I am experiencing *Q* now" at *t*, this introspective judgment is still distinct from the phenomenally conscious experience it is about—*provided that "Q" stands for a non-direct simple phenomenal concept*.

Thus, in order to stay on safe ground, the proponent of infallibility has to modify her claim once again. It makes sense that she restricts its scope to introspective judgments that involve direct phenomenal concepts. In this sense, Balog says:

[A] phenomenal concept may refer to a particular type of visual experience, say the experience typically caused by seeing red objects in ordinary light, etc.—call this type of experience “reddish”—by being constituted in part by a particular token of that type of experience. Then if I form the judgment I HAVE R where R is a direct phenomenal concept of reddish, my judgment cannot fail to be true. (Balog 2012, 27)<sup>26</sup>

Whether Balog’s claim is true depends on what is meant by “I have R.” A natural way of understanding this phrase is to say that one has R only if one visually experiences something red in the sense of being under the impression of facing a red object. Upon closer examination, however, it transpires that this *cannot* be the sense that Balog attaches to the phrase “having R”—for, under this interpretation, the judgment “I have R” seems to be fallible. In order to see the point, allow me to draw an analogy to linguistic quotation. When we put a word within linguistic quotation marks, it lacks its normal reference. Whereas the word “dog,” for example, refers to dogs, the word ““dog”” refers, not to dogs, but to the word “dog.” One might say, then, that linguistic quotation draws one’s attention away from the worldly item that is usually represented by the quoted word and toward the item that does the representing—namely, the word itself.

It seems plausible that something similar takes place in the case of mental quotation. Consider the visual experience of a ripe tomato. Putting this experience between mental quotes will draw your attention away from the worldly item that is usually represented by this experience, the ripe tomato. Instead, your attention will be shifted towards the item that does the representing—a multitude of phenomenal properties arranged in a certain order. Thus, as soon as your experience is mentally quoted it will lose a good deal of its intentionality: It changes from a representation of

---

<sup>26</sup> Balog uses capital letters to indicate that the relevant words stand for concepts in the language of thought.

an external material object to a representation of its (the representation's) phenomenal character.<sup>27</sup> Now, it seems to me that employing this latter kind of representation in thought—as it happens when you judge “I have R”—cannot count as a case of visually experiencing something red in the sense of being under the impression of facing a red object. Employing a representation of phenomenal redness does not mean being intentionally related to the color of the ripe tomato out there. Rather, it means apprehending a certain property of your experience in virtue of which the ripe tomato appears to be red, namely phenomenal redness. Apprehending phenomenal redness, in turn, does not amount to the impression of facing something red. Apprehending phenomenal redness is more akin to having a *sensation* such as an after-image or a phosphene, which likewise do not represent anything beyond themselves. Thus, the judgment “I have R”—understood as implying that you are under the impression of facing a red object—is simply false at the very moment you make it.

If these considerations are correct, then it is impossible to visually experience something as red *and* to employ a direct phenomenal concept of this very experience in thought at one and the same point in time. This impossibility claim may strike one as implausible. At least, it does not seem to us that we stop visually experiencing something as red when attending to the phenomenal property in virtue of which we experience something as red. However, I think that this is an illusion. Imagine that you are looking at a painting that fits into the conventional category of a portrait. There are

---

<sup>27</sup> An anonymous referee has pointed out that *if* we take the analogy to linguistic quotation seriously, the experience in quotation marks has no intentionality at all: “dog” in “dog” does not refer to anything; it is as little a signifying word as “nine” in “canine.” Thus, if experiences in mental quotes behave similarly to words in linguistic quotes, a mentally quoted experience does not have any representational power—it is just a representationally inert *part* of a more complex representational whole. The reason I recoil from this rather radical claim is that, as I said some paragraphs above, I am not sure whether we should take Balog’s talk of “mental quotes” literally. Linguistic quotation is not essentially tied to the use of quotation marks. There are other devices that may be used instead. Consider, for example, “*Dog* has three letters.” In this sentence, the word “dog” is quoted. Nevertheless, it does refer to something, namely to itself. Thus, there is some residual intentionality, so to speak. Similarly, mental quotation may not necessarily proceed by draping mental quotes around an experience. Thus, mental quotation may not deprive an experience of all of its intentionality.



at least two ways of looking at it. The first, and most natural, is to look at it as if it shows a person's face. The second, less natural, is to assume the stance that the French painter Maurice Denis once expressed as follows: "It should be remembered that a picture—before being a warhorse, a nude, or an anecdote of some sort—is essentially a flat surface covered with colors assembled in a certain order" (cf. Denis 1920, 1). When we look at the portrait from Denis's standpoint, the impression of being confronted with a person's face evaporates as all facial features dissolve into brush strokes, marks, and lines. It seems as if the portrait has vanished for us—there remains just a flat surface covered with paint. Thus, looking at the portrait as paint on canvas does not amount to the activity that is commonly called "viewing a portrait." Instead, it is more akin to the activity of viewing an *abstract painting*: We take the paint to represent itself, not anything *beyond* itself.

Now, the crucial point is that we cannot simultaneously assume both the normal attitude, whereby we are under the impression of being confronted with a human face, and the attitude described by Denis, whereby we are under the impression of being confronted with a flat surface covered with colors assembled in a certain way. However, we can easily switch between the two attitudes as we typically do when we study a painting in order to find out how the painter has achieved a certain realistic effect, say a reflection in an eye, by distributing the paint in a certain way. Thus, by switching attitudes over and over, it may eventually appear to us as if we are capable of simultaneously attending to the human face as it is represented by the portrait *and* to the brush strokes, marks, and lines that do the representing.

A similar story can be told in order to explain the impression that we do not stop visually experiencing something red when we employ a direct phenomenal concept of this very experience in thought. One moment, we attend to what our experience

represents—say, the color of a ripe tomato—, the next, we attend to that which does the representing—the phenomenal redness inherent to the experience. By switching attitudes over and over, it may eventually appear to us as if we are capable of simultaneously attending to what our experience represents and to the phenomenal properties in virtue of which our experience represents what it does.

It might be objected that all this does not show that it is *impossible* to simultaneously attend both to the content and the vehicle of one’s experience. However, even if such twofold attention were possible, this would not rule out cases in which—while judging “I have R”—one *exclusively* attends to the vehicle of one’s experience, thereby losing track of its content. For the sake of charitable reading of Balog, then, let us take the judgment “I have R” as *not* implying that the subject is under the impression of facing a red object. Rather, let us suppose that “I have R” is true even if the subject merely apprehends the phenomenal property of redness. *To be had*, in this sense, is the lowest common factor that indiscriminately characterizes any item that may be said to appear within, or to be present to, one’s mind: experiences, ideas, impressions, phenomenal properties, qualia, sensations, sense data, or what have you. It seems, then, that the following version of the infallibility claim might be immune to any objection—where the diamond symbolizes a current conscious experience and “\*◇\*” is a direct phenomenal concept that is generated by mentally quoting this experience:

Infallibility<sub>6</sub>: Necessarily, if *S* introspectively judges “I am having \*◇\* now” at *t*, *S*’s judgment is true.

However, recall that, according to Balog, a direct phenomenal concept that was generated by mentally quoting a current conscious experience might be understood in

two different ways. Depending on context, one might say that the concept in question refers to the *very token experience from which it is built*, or one might say that it refers to *a certain kind of experience*.<sup>28</sup> Rendered in the first way, “I am having \*◇\* now” is true if and only if the subject has the very token that has been put between mental quotes. By contrast, according to the second interpretation, “I am having \*◇\* now” is true if and only if the subject has a token of the kind to which “\*◇\*” refers. Now, I think that the second interpretation leaves room for doubts whether *Infallibility*<sub>6</sub> really holds.

Suppose that you have the visual experience of a ripe tomato as illustrated in figure 1.



Figure 1

Suppose furthermore that you take this experience and put it between mental quotation marks. What results, says Balog, is a concept that refers “to a type of phenomenal experience that the token exemplifies” (Balog 2012, 34). Notice, however, that your visual experience of a ripe tomato exemplifies *more than one phenomenal type*. It might be said, for example, that your visual experience of a ripe tomato is of the same phenomenal type as the visual experience you had yesterday when you looked at a red patch, as illustrated in figure 2.



---

<sup>28</sup> Cf. Balog (2012, 34). I have quoted the passage above.

Figure 2

Similarly, it might be said that your visual experience of a ripe tomato is of the same phenomenal type as the visual experience illustrated by figure 3.



Figure 3

Both experiences exemplify a phenomenal property that may be called *phenomenal tomato-shapedness*. Both experiences are phenomenal-tomato-shapedness-experiences, so to speak. Briefly, the experience that is put between mental quotes does not *determine* the kind of phenomenal experience to which the resulting direct phenomenal concept refers.<sup>29</sup> Thus, in order to produce a direct phenomenal concept with a determinate reference, there has to be something like an act of abstraction: one must attend to one of the many phenomenal properties that are exemplified by one's experience and ignore the rest.<sup>30</sup>

Seen in this light, mental quotation is an instance of what Keith Lehrer calls *exemplarization*: One uses a particular *F* in order to represent a certain kind of things,

---

<sup>29</sup> Balog explicitly admits this point. Cf. Balog (2012, 37 f.).

<sup>30</sup> Balog's own answer to the question, "What determines the type a phenomenal concept refers to, if the token experience that constitutes it doesn't, or at least doesn't fully, determine it?" refers to the conceptual role of the phenomenal concept in question: "A particular token of the concept **\*experience e\*** refers to, e.g., reddish experiences if, were I confronted with any kind of reddish experience, I would judge "same kind of experience." It refers to dark-reddish experiences if, were I confronted with dark-reddish experiences, I would judge "same kind of experience" but not so when I am confronted with light reddish experiences, etc." (Balog 2012, 37 f.) I do not claim that Balog's answer is wrong. However, Balog's answer is incomplete since she does not tell us how a given phenomenal concept *acquires* its characteristic conceptual role. In my view, it seems most plausible to say that phenomenal concepts acquire their characteristic conceptual roles through a process of abstraction.

namely things that exemplify the property of *being F*.<sup>31</sup> One may use, for example, a particular cat to represent a certain kind of creature, say cats. One thereby abstracts from all those properties of the particular cat that are not entailed by the property of *being a cat*. The same applies to the formation of a direct phenomenal concept of, say, phenomenal redness. A particular experience of something red is “exemplarized,” as Lehrer says, and thereby “becomes conceptual” (Lehrer 2002, 423). The subject uses a particular experience to represent a certain kind of experience—experiences that instantiate phenomenal redness—by abstracting from all those properties of the particular experience that are not entailed by the phenomenal property of redness.

Now, as Lehrer remarks, “psychology may go awry in the activity of exemplarization, and I might, in fact, use the exemplarized state to represent a class of states that does not include the exemplar itself” (Lehrer 2002, 428). Consider, again, the visual experience of a ripe tomato as illustrated in figure 1. A token of this experience is an exemplar of a certain kind of experience, for instance, experiences that instantiate the phenomenal property of tomato-shapedness. Suppose then that you attend to phenomenal tomato-shapedness as it is exemplified by your experience and abstract from all other phenomenal properties that might also be salient. Due to some kind of cognitive pushiness, however, you automatically correct for the irregularities of the tomato-shape. As a result, you do not form a conception of phenomenal tomato-shapedness, but say, of phenomenal perfect-circularity. Therefore, although your current visual experience is part of the direct phenomenal concept you form on that occasion, it does not fit that concept—for it does not instantiate phenomenal perfect-circularity but only phenomenal tomato-shapedness.<sup>32</sup> If you judged, in this situation, “I am having

---

<sup>31</sup> Cf. Lehrer (1997; 2002; 2006).

<sup>32</sup> This example is supposed to be strictly analogous to a case that is mentioned by Lehrer (2006, 412 f.) in order to illustrate his claim that an exemplar might fail to be an instance of the kind it represents.

\*◇\* now”—where the diamond stands for your current visual experience of a ripe tomato and “\*◇\*” for your direct phenomenal concept of phenomenal perfect-circularity—you would be wrong.<sup>33</sup>

#### 4. Infallibility<sub>7</sub> and the Role of Acquaintance

If the considerations of the last few paragraphs are correct, then judgments of the form “I am having \*◇\* now” are fallible—provided that we take “\*◇\*” as referring to a phenomenal kind. Thus, the proponent of infallibility would be well advised to restrict the interpretation of “\*◇\*” to the one according to which “\*◇\*” refers to the very token experience from which it is built. This interpretation does not presuppose any act of exemplarization. Hence, mistakes in the process of exemplarization cannot occur:

Infallibility<sub>7</sub>: Necessarily, if *S* introspectively judges “I am having \*◇\* now” at *t*—where “\*◇\*” is a direct phenomenal concept that refers to the very token experience from which it is built—, *S*’s judgment is true.

I admit that I cannot find any counterexamples to *Infallibility*<sub>7</sub>. Yet, I still think that it does not do the trick since *Infallibility*<sub>7</sub> does not show that acquaintance explains infallibility. Recall that proponents of infallibility usually do not argue for infallibility for its own sake. Instead, they are interested in establishing a certain claim about the relation each of us bears to his or her own current experiences. The claim is that there is a special relation between ourselves and our own current experiences—a relation that

---

<sup>33</sup> I admit that this case is, given the actual psychology of normal human beings, extremely unlikely. However, I would not vouch for the claim that, even given the actual psychology of normal human beings, this case is *impossible*.

we do not bear to any other item, neither to our past experiences, nor to our bodily states, nor to the experiences of others, nor to material things in the external world. In a word, we are *acquainted* with our current experiences. Moreover, it is said that acquaintance<sup>34</sup> makes possible a “certain sort of lucid understanding” (Chalmers 2003, 248) of those experiences. It suggests itself, then, that acquaintance *explains* why particular judgments about the phenomenal properties of one’s own current experiences are infallible: They are infallible because they are the results of a kind of inner intuition that presents its objects as they really are. Engaging in this kind of intuition cannot possibly lead to false judgments.

As I indicated at the beginning of this paper, this traditional view is currently enjoying a revival. For example, David Chalmers, immediately after he has developed and defended a version of the infallibility thesis in the spirit of *Infallibility*<sub>7</sub>, writes:

At this point it is natural to ask: if we can form this special class of incorrigible, distinctively constituted beliefs where phenomenal states and properties are concerned, why cannot we do so where other states and properties are concerned? ... [I]t is natural to suggest that ... we bear a special relation to the phenomenal properties instantiated in our experience: a relation that we do not bear to the other instantiated properties in question, and a relation that is required in order to form a direct concept of a property in the manner described. This relation would seem to be a peculiarly intimate one, made possible by the fact that experiences lie at the heart of the mind rather than standing at a distance from it ... We might call this relation *acquaintance*. (Chalmers 2003, 246–248)<sup>35</sup>

---

<sup>34</sup> Alternatively, Horgan & Kriegel (2007) speak of “inner awareness.” A third notion for the same phenomenon might be “direct awareness.”

<sup>35</sup> Unlike me, who—following Balog—conceives of experiences as mental individuals that exemplify phenomenal properties, Chalmers takes experiences for instantiations of phenomenal properties. Thus, Chalmers construes acquaintance as a relation to phenomenal properties, whereas I—following Balog—take acquaintance as a relation to experiences. Cf. FN 1, where I explain these two different ways of speaking about the mental in more detail.

A natural way to understand Chalmers is as follows: (i) there is a particular type of judgments about one's own current phenomenal experiences that are infallible; (ii) one could not have this particular type of infallible judgment unless one could form direct phenomenal concepts; (iii) one could not form direct phenomenal concepts unless one was acquainted with some of one's own current phenomenal experiences; (iv) thus, acquaintance explains why judgments of the particular type are infallible.

I do not find this argument very convincing. However, my point is not that there is no particular type of judgments about one's own current phenomenal experiences that are infallible, or that one could not have this particular type of infallible judgment unless one could form direct phenomenal concepts. On the contrary, I believe that there is such a particular type, namely "I am having \*◇\* now" (on the relevant reading). Thus, it cannot be denied that, in order to make such a judgment, one has to form a direct phenomenal concept. Nor do I deny that there is a relation between subjects and experiences—call it "acquaintance" if you like—that makes possible the formation of direct phenomenal concepts. However, I deny that acquaintance *explains* why judgments of the type "I am having \*◇\* now" (on the relevant reading) are infallible.<sup>36</sup>

Let us consider an analogous case from linguistic quotation again. Take, for example, utterances of the type "I am uttering »shoo-be-doo-be-doo« now"—where the French quotes stand for Reichenbach's token quotes.<sup>37</sup> So the token of "»shoo-be-doo-be-doo«" that appears in an utterance of "I am uttering »shoo-be-doo-be-doo« now" does not refer to the phrase "shoo-be-doo-be-doo" considered as a type but to the very token of "shoo-be-doo-be-doo" as it appears in the utterance. Thus, any particular utterance of "I am uttering »shoo-be-doo-be-doo« now" will turn out to be true.

---

<sup>36</sup> I understand the considerations that follow as a development of some remarks by Shoemaker (1996, 53) and Schwitzgebel (2008, 261) who both account for the infallibility of certain introspective judgments, not in terms of an alleged privileged epistemic relation we bear to aspects of our own mind, but in terms of the self-verifying character of those judgments.

<sup>37</sup> Cf. Reichenbach (1947, 284). Reichenbach uses little arrows instead of French quotes.



Utterances of the type “I am uttering »shoo-be-doo-be-doo« now,” then, have certain similarities to utterances of sentences that traditionally are supposed to be *contingently a priori*. As in the case of “I am here now,” for example, what one asserts when one utters “I am uttering »shoo-be-doo-be-doo« now” is not true at every possible world. For there are possible worlds in which the person to which “I” refers is silent (or utters different words) at the time to which “now” refers. However, it seems that we know in advance that, if somebody were to utter “I am uttering »shoo-be-doo-be-doo« now” he or she could not avoid speaking the truth.<sup>38</sup> Now, essentially the same seems to apply to judgments of the type “I am having \*◇\* now”: These judgments likewise do not convey any necessary truths but are just true on every possible occasion they are made. Thus, it might be appropriate to say that utterances of the type “I am uttering »shoo-be-doo-be-doo« now” and judgments of the type “I am having \*◇\* now” share *the same kind of infallibility*.

Now, it seems obvious that one could not make an utterance of “I am uttering »shoo-be-doo-be-doo« now” unless one could form quoting expressions. It seems equally obvious that one could not form, for example, the quoting expression “»shoo-be-doo-be-doo«” unless one bore a certain relation to the quoted phrase “shoo-be-doo-be-doo.” One might describe this relation by saying that the utterer has the phrase “shoo-be-doo-be-doo” *at his or her disposal ready for being quoted*. This is just the relation we bear to all, and only those, phrases such that we are on the point of putting

---

<sup>38</sup> In the literature, there is some disagreement about the question whether there might be false utterances of sentences like “I am here now,” given that “I” refers to the utterer, “here” refers to the place where the utterance is made, and “now” refers to the time at which the utterance is made. For example, suppose that Jill has recorded “Hi, I am here now, but I don’t want to pick up the phone. So please leave a message” on her answering machine. Suppose furthermore that she puts her phone on answering machine only if she is not at home. Now, imagine that you call her but, since she is not at home, you only reach the answering machine. This might be regarded as a case where there is a *wrong* utterance of “I am here now”—namely, the recording you hear on that occasion. For the utterer—Jill—is not at the place where the utterance is made—Jill’s home—at the time when the utterance is made, i.e., when you call. (Cf. Predelli (2005, 43–46).) However, I take it that my example (“I am uttering »shoo-be-doo-be-doo« now”) is immune against objections of that kind. At least, I am not able to imagine a situation in which the sentence “I am uttering »shoo-be-doo-be-doo« now” is uttered falsely.

them between quotes. Echoing the words of Chalmers, one might even say that this relation is an “intimate one, made possible by the fact that language lies at the heart of the mind rather than standing at a distance from it.”

However, it does not seem right to me to say that this relation—the having-at-one’s-disposal-ready-for-being-quoted—explains why utterances of “I am uttering »shoo-be-doo-be-doo« now” feature the kind of infallibility they do. In my view, the infallibility of these utterances stems from the fact that *they cannot, even not possibly, exist if the state of affairs that makes them true does not obtain*: If one did not utter a token of “shoo-be-doo-be-doo” at time  $t$  one could not make the utterance “I am uttering »shoo-be-doo-be-doo« now” at  $t$ . Let us call this phenomenon *self-verification* and explain it as follows:

For all truth-apt representations  $x$ :  $x$  is self-verifying if and only if, necessarily,  $x$  would not exist if  $x$  were not true.

The reason why utterances of “I am uttering »shoo-be-doo-be-doo« now” are infallible, then, is that these utterances are self-verifying—not that the subject has the phrase “shoo-be-doo-be-doo” at her disposal ready for being quoted. Having the phrase “shoo-be-doo-be-doo” at one’s disposal ready for being quoted is, to be sure, a necessary condition for *forming* the quoting expression “»shoo-be-doo-be-doo«” and, hence, a necessary condition for *uttering* “I am uttering »shoo-be-doo-be-doo« now.” But the fact that the phrase “shoo-be-doo-be-doo” is at one’s disposal ready for being quoted *does not explain why it is impossible to make an utterance of “I am uttering »shoo-be-doo-be-doo« now” that turns out to be false*.

Notice that, on the one hand, there are many other utterances that employ the quoting expression “»shoo-be-doo-be-doo«,” and thus likewise rely upon the fact that

the subject has the phrase “shoo-be-doo-be-doo” at her disposal, ready for being quoted, that are *not* infallibly true. Consider, for example, “I do not utter »Shoo-be-doo-be-doo« now.” On the other hand, there are utterances that are in the same sense infallible as “I am uttering »shoo-be-doo-be-doo« now” that do *not* employ any quoting expression. Consider, for example, “I am uttering something now.” This utterance likewise does not convey any necessary truth. For there might be possible worlds in which the person to which “I” refers does not utter anything at the time to which “now” refers. However, the sentence seems to be true on every possible occasion it is uttered. Yet uttering the sentence, “I am uttering something now” does not presuppose that one bears the relation of having-at-one’s-disposal-ready-for-being-quoted to any expression the sentence contains. Having a certain phrase at one’s disposal ready for being quoted, then, is neither necessary nor sufficient for the kind of infallibility that an utterance of “I am uttering »shoo-be-doo-be-doo« now” has. In contrast, the feature of being self-verifying seems to be sufficient for the kind of infallibility in question. At least I cannot imagine a sentence that might be self-verifying, but not true on every possible occasion it is uttered. So I conclude that it is this latter feature—their self-verifying character—that explains why utterances of “I am uttering »shoo-be-doo-be-doo« now” are guaranteed to be true.

Now, I think that essentially the same applies to judgments of the type “I am having \*◇\* now.” First, note that “acquaintance” is introduced by Chalmers as a theoretical notion that refers to “that relation between subjects and properties that makes possible the formation of direct phenomenal concepts” (Chalmers 2002, 248).<sup>39</sup> According to this understanding of acquaintance, the claim that while judging “I am

---

<sup>39</sup> See also Balog (2009, 292): “When we deploy phenomenal concepts introspectively for some phenomenally conscious experience as it occurs, we are said to be *acquainted* with our own conscious experiences” (original emphasis).

having \*◇\* now” one is acquainted with the experience symbolized by the diamond is just trivially true. However, the fact that I am acquainted with this experience does not *explain* why my judgment features the kind of infallibility it does.

Recall that there are other introspective judgments that employ direct phenomenal concepts, and thus likewise rely upon acquaintance, that are *not* infallibly true in the sense that “I am having \*◇\* now” is. Consider, for example, “I am not having \*◇\* now.” On the other hand, there are introspective judgments that are in the same sense infallible as “I am having \*◇\* now” that do not employ any direct phenomenal concepts. Consider, for example, judgments of the type “I am thinking that *p* now”—where “to think that *p*” is supposed to mean “to entertain, in whatever mode, the thought that *p*.” Judgments of this type likewise do not convey any necessary truth. For there might be possible worlds in which the person to whom “I” refers collapses unconscious (or is thinking a different thought) at the time to which “now” refers. However, the mere fact that one judges “I am thinking that *p* now” does not imply that one stands in an acquaintance relation to something.<sup>40</sup> Even a phenomenal zombie might be able to judge that he or she is thinking that *p* now. Acquaintance, then, is neither necessary nor sufficient for making an introspective judgment that is infallibly true in the sense that “I am having \*◇\* now” is. Thus, it seems that acquaintance is just an arbitrary factor in the production of infallibility: it might play a role in bringing judgments of the type “I am having \*◇\* now” into existence, but it does not explain

---

<sup>40</sup> This is a contentious claim. Advocates of cognitive phenomenology will deny it, arguing that the propositional contents of our conscious thoughts have phenomenal properties, and that we introspect these contents by identifying their phenomenal properties. Thus, introspective awareness of the contents of our own conscious thoughts necessarily involves some sort of acquaintance (cf. Pitt (2004; 2011)). I am skeptical about the whole idea of cognitive phenomenology. However, for reasons of space, I cannot argue against it here. For some good arguments against cognitive phenomenology, see Carruthers & Veillet (2011) and Tye & Wright (2011). Thanks to an anonymous referee for bringing this issue to my attention.

why those judgments are guaranteed to be true on every possible occasion they are made.

My positive view is that the infallibility of “I am having \*◇\* now” is due to the fact that this judgment could not even exist if the state of affairs that makes it true did not obtain: If I did not sense a token of “◇” at time *t*, I could not judge “I am having \*◇\* now” at *t*. It is *this*, the phenomenon of self-verification, that explains why my judgment is guaranteed to be true on every occasion it is made—it is *not* the fact that I am acquainted with the experience I mentally quote. Being acquainted with that experience is, certainly, a necessary condition for *forming* the concept “\*◇\*” and, hence, a necessary condition for *judging* “I am having \*◇\* now.” Nevertheless, acquaintance does not provide an answer to the question why those judgments are guaranteed to be true. For, as we have seen, there are both judgments that are infallible although they are not based on acquaintance and judgments that might be based on acquaintance, but are not infallible at all.

Before I present my conclusion, allow me to address a final doubt about my argument.<sup>41</sup> If my argument is sound, one might say, then acquaintance does not play any explanatory role in connection with judgments of the “I am having \*◇\* now”-form, except that it is necessary for their formation. However, this seems wrong. Suppose that I am undergoing a particular experience at time *t*, say the experience of a ripe tomato. Given the required cognitive capacity, it is possible for me to form a direct phenomenal concept that involves this very experience and deploy it in judgments about my current mental state. There are many such judgments that I could make. Suppose, however, that, at *t*, I am forced to choose between “I am having \*◇\* now” and “I am not having \*◇\* now”—where the diamond symbolizes my current experience of a ripe tomato. Of

---

<sup>41</sup> I am grateful to an anonymous referee for drawing this objection to my attention.

course, I will choose the first and not the second, since I *know* at *t* that “I am having \*◇\* now” is true and “I am not having \*◇\* now” is false. At least, I am in a position to know which judgment is right and which is wrong. But how is this possible? The best explanation, it seems, is that I know this because I am *acquainted* with the experience that I undergo at *t*. Thus, acquaintance does play a significant role in connection with judgments of the “I am having \*◇\* now”-form, besides being necessary for their formation: Acquaintance explains why I am in a position to know why judgments of the “I am having \*◇\* now”-form are true and their corresponding negations are false.

I do not find this argument very convincing. First, it confounds two questions that are inherently separate: the question of why judgments of the “I am having \*◇\* now”-form cannot be false, and the question of why *we are in a position to know* that judgments of the “I am having \*◇\* now”-form are true and their corresponding negations are false. It is the first question that I have tried to answer in this paper, not the second one. It may be that acquaintance does play a significant role in answering the second question, but this does not imply that it plays a similar role in answering the first. Second, there is an alternative explanation of why we are in a position to know why judgments of the “I am having \*◇\* now”-form are true and their corresponding negations are false: The reason might be simply that one is *a priori* justified in believing (i) that judgments of the “I am having \*◇\* now”-form are self-verifying, (ii) that self-verifying judgments are true on every occasion they are made, and (iii) that negations of self-verifying judgments are false on every occasion they are made. Thus, the appeal to acquaintance seems to be dispensable even in connection with the question of why it is that we are in a position to know that judgments of the “I am having \*◇\* now”-form are true and their corresponding negations false.

## 5. Conclusion

The aim of this paper was to criticize the view that there is a close relationship between the infallibility of certain introspective judgments about one's own current experiences and the fact that we are acquainted with those experiences. The view in question starts out from the fact that there are particular judgments about the phenomenal properties of one's own current mental states that are infallibly true. The alleged explanation is that one bears a special relation to one's current phenomenal experiences—namely, acquaintance (or inner awareness). While being acquainted with (or internally aware of) an experience, the story goes, one cannot help apprehending the experience as it really is.

The first thing I have done in this paper was to reconstruct the reasons that led philosophers to the view that, if any, then only judgments about the phenomenal character of one's own current experiences are infallible. Then I asked *exactly which* judgments about the phenomenal character of one's own current experiences are infallible. It emerged that the answer to this question is by no means obvious. Finally, however, I arrived at judgments of the “I am having \*◇\* now”-form— where “\*◇\*” is a direct phenomenal concept formed by mentally quoting an experience one is currently undergoing.

The second thing I have done in this paper was to attack the view that judgments of the “I am having \*◇\* now”-form are infallible because subjects are acquainted with their current experiences. I criticized this view based on the observation that we might encounter acquaintance where there is no infallibility, but also might encounter infallibility where there is no acquaintance. Thus, I concluded that acquaintance is

expendable when it comes to an explanation of why judgments of the “I am having \*◇\* now”-form are infallible. I positively argued that the infallibility of such judgments is due to the fact that they could not possibly exist if they were not true. In a sense then, the infallibility of judgments of the “I am having \*◇\* now”-form is not the fruit of the peculiar relation each of us bears to his or her conscious experiences but of a peculiar interplay between the truth conditions and the existence conditions of those judgments.\*

#### REFERENCES

- ARMSTRONG, D. M. 1963, “Is Introspective Knowledge Incorrigible?”, *Philosophical Review* 72, pp. 417–432.
- ARMSTRONG, D. M. 1968, *A Materialist Theory of the Mind*, London: Routledge.
- BALOG, K. 2009, “Phenomenal Concepts”, in: B. McLaughlin, A. Beckermann and S. Walter (eds.), *Oxford Handbook in the Philosophy of Mind*, Oxford: Oxford University Press, pp. 292–312.
- BALOG, K. 2012, “Acquaintance and the Mind-Body Problem”, in: S. Gozzano and C. Hill (eds.), *New Perspectives on Type Identity. The Mental and the Physical*, Cambridge: Cambridge University Press, pp. 16–42.
- BAR-ON, D. 2004, *Speaking My Mind*, Oxford: Oxford University Press.
- BYRNE, A. 2002, “Don’t PANIC: Tye’s Intentionalist Theory of Consciousness”, *A Field Guide to the Philosophy of Mind* symposium on Tye’s *Consciousness, Color, and Content*.

---

\* First of all, I am greatly indebted to Thomas Spitzley without whose support the research that led to this paper would not have been possible. He also gave me the opportunity to present a previous draft to his graduate seminar at the University of Duisburg-Essen in June 2013. Thanks to all participants for their extremely valuable comments. Moreover, I would like to thank Tobias Steinig for all of his efforts: he not only proofread the entire manuscript, but also did extensive bibliographical research. Special thanks are due to the anonymous referees and editors of this journal whose constructive criticism greatly helped me to clarify my argument. Work on this paper was supported by research grant BA 2269/2-1 from the German Research Foundation (DFG).



URL=<[http://host.uniroma3.it/progetti/kant/field/tyesymp\\_byrne.htm](http://host.uniroma3.it/progetti/kant/field/tyesymp_byrne.htm)> (last accessed January 9, 2015).

CARRUTHERS, P. 2011, “The Case against Cognitive Phenomenology”, in: T. Bayne and M. Montague (eds.), *Cognitive Phenomenology*, Oxford: Oxford University Press, pp. 35–55.

CHALMERS, D. 2003, “The Content and Epistemology of Phenomenal Belief”, in: Q. Smith and A. Jokic (eds.), *Consciousness: New Philosophical Perspectives*, Oxford: Clarendon, pp. 220–272.

CHISHOLM, R. 1942, “The Problem of the Speckled Hen”, *Mind* 51, pp. 368–373.

DENIS, M. 1920, *Théories 1890–1910: Du Symbolisme et de Gauguin vers un Nouvel Ordre Classique*, Paris: L. Rouart et J. Watelin.

FISH, W. 2010, *Philosophy of Perception. A Contemporary Introduction*, London: Routledge.

GERTLER, B. 2001, “Introspecting Phenomenal States”, *Philosophy and Phenomenological Research* 63, pp. 305–328.

GERTLER, B. 2011, *Self-Knowledge*, London: Routledge.

GERTLER, B. 2012, “Renewed Acquaintance”, in: D. Smithies and D. Stoljar (eds.), *Introspection and Consciousness*, Oxford: Oxford University Press, pp. 93–127.

HORGAN, T. 2012, “Introspection about Phenomenal Consciousness: Running the Gamut from Infallibility to Impotence”, in: D. Smithies and D. Stoljar (eds.), *Introspection and Consciousness*, Oxford: Oxford University Press, pp. 405–422.

HORGAN, T. and KRIEGEL, U. 2007, “Phenomenal Epistemology: What Is Consciousness That We May Know It So Well?”, *Philosophical Issues* 17, pp. 123–144.

JACKSON, F. 1973, “Is There a Good Argument Against the Incorrigenibility Thesis?”, *Australasian Journal of Philosophy* 51, pp. 51–62.

KIDD, C. 2011, “Phenomenal Consciousness with Infallible Self-Representation”, *Philosophical Studies* 152, pp. 361–383.

LEHRER, K. 1997, *Self-Trust*, Oxford: Oxford University Press.

LEHRER, K. 2002, “Self-presentation, Representation and the Self”, *Philosophy and Phenomenological*

*Research* 64, pp. 412–430.

LEHRER, K. 2006, “Consciousness, Representation, and Knowledge”, in: U. Kriegel and K. Williford (eds.), *Self-Representational Approaches to Consciousness*, Cambridge, MA: MIT Press, pp. 409–419.

LEWIS, C. I. 1946, *An Analysis of Knowledge and Valuation*, La Salle: Open Court.

LYCAN, W. 1996, *Consciousness and Experience*, Cambridge, MA: MIT Press.

NEANDER, K. 1998, “The Division of Phenomenal Labour: A Problem for Representational Theories of Consciousness”, *Philosophical Perspectives* 12, pp. 411–434.

PAPINEAU, D. 2002, *Thinking about Consciousness*, Oxford: Oxford University Press.

PAPINEAU, D. 2007, “Phenomenal and Perceptual Concepts”, in: T. Alter and S. Walter (eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, Oxford: Oxford University Press, pp. 111–144.

PITT, D. 2004, “The Phenomenology of Cognition, Or, What Is It Like to Think That P?”, *Philosophy and Phenomenological Research* 69, pp. 1–36.

PITT, D. 2011, “Introspection, Phenomenality, and the Availability of Intentional Content”, in: T. Bayne and M. Montague (eds.), *Cognitive Phenomenology*, Oxford: Oxford University Press, pp. 142–172.

PREDELLI, S. 2005, *Contexts: Meaning, Truth, and the Use of Language*, Oxford: Oxford University Press.

RAFF, C. 1966, “Introspection and Incorrigeability”, *Philosophy and Phenomenological Research* 27, pp. 69–73.

REICHENBACH, H. 1947, *Elements of Symbolic Logic*, New York: Macmillan.

RYLE, G. 1949, *The Concept of Mind*, London: Hutchinson’s.

SCHWITZGEBEL, E. 2008, “The Unreliability of Naive Introspection”, *Philosophical Review* 117, pp. 245–273.

- SCHWITZGEBEL, E. 2012, “Self-Ignorance”, in: J. Liu and J. Perry (eds.), *Consciousness and the Self: New Essays*, Cambridge: Cambridge University Press, pp. 184–197.
- SHOEMAKER, S. 1994, “Self-Knowledge and Inner Sense II: The Broad Perceptual Model”, *Philosophy and Phenomenological Research* 54, pp. 271–290.
- SHOEMAKER, S. 1996, “First Person Access”, in: S. Shoemaker, *The First Person Perspective and Other Essays*, Cambridge: Cambridge University Press, pp. 50–73.
- SMART, J. J. C. 1962, “Brain Processes and Incorrigeability”, *Australasian Journal of Philosophy* 40, pp. 68–70.
- SOSA, E. 2003, “Privileged Access”, in: Q. Smith and A. Jokic (eds.), *Consciousness: New Philosophical Perspectives*, Oxford: Clarendon, pp. 273–292.
- TYE, M. and WRIGHT, B. 2011, “Is There a Phenomenology of Thought?”, in: T. Bayne and M. Montague (eds.), *Cognitive Phenomenology*, Oxford: Oxford University Press, pp. 326–343.
- WEISBERG, J. 2008, “Same Old, Same Old: The Same-Order Representation Theory of Consciousness and the Division of Phenomenal Labor”, *Synthese* 160, pp. 161–181.