

## IS THERE ANYTHING TO THE AUTHORITY THESIS?

Wolfgang Barz

Johann Wolfgang Goethe University Frankfurt am Main

barz@em.uni-frankfurt.de

*Abstract.* Many philosophical theories of self-knowledge can be understood as attempts to explain why self-ascriptions enjoy a certain kind of authority that other-ascriptions lack (the Authority Thesis). The aim of this paper is not to expand the stock of existing explanations but to ask whether the Authority Thesis can be adequately specified. To this end, I identify three requirements that must be met by any satisfactory specification. I conclude that the search for an adequate specification of the Authority Thesis leads to a dilemma: it either yields an interpretation under which the thesis is philosophically interesting but false, or it produces an interpretation under which the thesis is actually true but of minor philosophical interest.

## I. INTRODUCTION

Philosophical discussions about the utterances we use to attribute current mental states to ourselves—commonly called ‘self-ascriptions,’ ‘first-person reports,’ or ‘avowals’—usually follow a simple script. Self-ascriptions are said to be treated with a special kind of deference with which the utterances we use to attribute current mental states to *other people*—termed ‘other-ascriptions’—are not treated. This difference seems puzzling because corresponding pairs of self- and other-ascriptions, such as ‘I am in pain’ and ‘She is in pain,’ do not differ with respect to the state of affairs they express: both ascriptions concern the same contingent matter of fact—a particular person’s mental state. How, then, is the special status of self-ascriptions to be explained?

Undoubtedly, this way of framing the matter is relatively vague. However, the important point is that a similar way of exposing the problem is pervasive in the literature. The introductory passages in Dorit Bar-On’s *Speaking My Mind* are an especially clear example:

My main task in this book is to provide an account of the unique status assigned to avowals .... I take this task to be of considerable philosophical interest. For, even if avowals do not enjoy absolute security, their relatively high degree of security as compared with all other empirical statements should make us wonder. Why should my own pronouncements regarding my present states of mind ... carry so much more weight than anyone else’s pronouncements on the same matters? (Bar-On 2004: 10)

Crispin Wright (1998: 106) provides another example when he writes that “The cardinal problem of self-knowledge is that of explaining *why* avowals display the marks they do

... what is the source of the special authority carried by their verdicts [i.e., subjects' avowals]?"<sup>1</sup>

As is typical in philosophy, there are almost as many answers to these questions as there are authors who address this issue. However, this paper does not aim to increase the stock of existing answers. Instead, I will ask what, exactly, is the phenomenon that Bar-On, Wright, and many other philosophers try to explain. Typically, philosophers use metaphors to describe the phenomenon they are after: Bar-On, for example, states that self-ascriptions "carry more weight" than other-ascriptions, and Wright speaks of an "authority" that is "carried" by self-ascriptions. Though these formulations may hint at the phenomenon for which an explanation is sought, they are amenable to a wide range of different interpretations. Thus, my discussion will question whether there is a satisfactory interpretation of what I will call the Authority Thesis:

Authority Thesis                  Self-ascriptions enjoy a kind of authority that other-ascriptions lack.

Admittedly, the Authority Thesis is extremely vague. However, that is precisely its purpose: the Authority Thesis is *intended* to be a provisional formulation needing specification. Only when an adequate specification is found will we be equipped to know exactly which phenomenon Bar-On, Wright, and other philosophers attempt to explain. Thus, this paper will proceed as follows. In Section 2, I present three requirements that any adequate specification of the Authority Thesis must meet. In Sections 3, 4, and 5, I discuss several initially promising interpretations of the Authority Thesis, and show that none of them satisfies all three requirements. This does not mean that the Authority Thesis is *false* under every interpretation. On the contrary, at least one interpretation will emerge under which the Authority Thesis is true. Nonetheless,

that interpretation is inadequate since it renders the Authority Thesis uninteresting in philosophical respects. Thus, I conclude that the search for an adequate specification of the Authority Thesis leads to a dilemma: it either yields an interpretation under which the thesis is philosophically interesting but false, or it produces an interpretation under which the thesis is actually true but of minor philosophical interest.

First, I want to set aside a misunderstanding. Philosophers use the term ‘authority’ in different ways. For example, some authors use it to refer to the alleged fact that my *knowledge* about my present mental states is somehow superior to the knowledge anyone else might have about them: I seem to know best what I currently believe, want, experience, and feel.<sup>2</sup> Thus, one might be inclined to consider the Authority Thesis tantamount to some privileged access thesis.<sup>3</sup> However, this inclination would draw an erroneous parallel. First, privileged access theses concern certain types of *beliefs*—namely, beliefs about one’s present mental states—whereas the Authority Thesis concerns *utterances*—namely, the utterances one uses to attribute current mental states to oneself. Moreover, privileged access theses contrast the epistemic properties of said beliefs with the epistemic properties of beliefs about others’ mental states, whereas the Authority Thesis contrasts the impact of said utterances on their addressees with the respective impact of utterances one uses to attribute current mental states to others. ‘Authority’ in the sense of the Authority Thesis, then, is not something that subjects enjoy about their own current mental states but rather something that verbal utterances enjoy in the ear of their hearers, so to speak.<sup>4</sup> I do not deny that the authority of self-ascriptions might be *explained* in terms of privileged access. My point is simply that the notion of authority, as it appears in the Authority Thesis, concerns a pragmatic property of utterances, not an epistemic property of beliefs. In a sense, then, this paper’s scope is rather narrow compared with the research

commonly practiced under the heading of “self-knowledge”: I will focus almost exclusively on whether there is a philosophically interesting way to capture the impact of self-ascriptions on their hearers—I will *not* be concerned with whether beliefs about one’s own present mental states have an epistemic status that other empirical beliefs lack.

## **II. THREE REQUIREMENTS FOR ANY SATISFACTORY SPECIFICATION OF THE AUTHORITY THESIS**

To launch my criticism, I first formulate three basic requirements for any satisfactory specification of the Authority Thesis. First, any satisfactory specification must delineate the alleged authority as a feature that at least a sufficiently large number of *self-ascriptions have*. This, I think, is rather trivial and does not require additional explanation. The same holds for the second requirement: any satisfactory specification must delineate the alleged authority as a feature that *the third-person counterparts of self-ascriptions—in short, “other-ascriptions”—lack*. The third requirement, however, is not so trivial: any satisfactory specification must delineate the alleged authority as a feature that *gives rise to a kind of puzzlement distinctive of philosophical problems*. Because including this requirement is not a matter of course, allow me to elucidate further.

Tradition has it that philosophy begins in wonder, astonishment, or perplexity. Typically, philosophical theorizing is provoked by certain remarkable features of reality; features of which we seem unable to make sense, even with the help of empirical science. Thinking about these features can easily lead to the state of puzzlement Wittgenstein once characterized by the phrase, ‘I don’t know my way about’ (PI §123). Very often, this kind of puzzlement can be represented as an apory,

that is, a group of statements that are individually plausible but collectively inconsistent. A prime example of such an apory is the mind-body problem. Consider the following three statements:

- (1) The mental and the physical are distinct.
- (2) The mental and the physical causally interact.
- (3) The physical is causally closed.<sup>5</sup>

Each of these statements is intuitively compelling: (1) is part of our common sense understanding of the mental as opposed to the physical; (2) might be made plausible by examples of bodily movements (physical phenomena) that result from decisions (mental phenomena), such as raising one's arm during a political convention; finally, (3) formulates a methodological principle of natural science. Yet the truth of any two of these statements implies the falsity of the remaining third. Moreover, it is characteristic of these kinds of problems that they persist despite the progress of natural science. Thus, it seems that something more is needed to solve them. A philosophical problem, then, might be characterized as a 'clash of intuitions' that do not seem reconcilable by empirical means alone. The philosopher's task is to approach, and maybe solve, these problems from the armchair: by rational argumentation, flanked by conceptual analysis, thought experiments, and other non-empirical methods, such as imagining counterfactual situations.<sup>6</sup>

This metaphilosophical view has some important repercussions for our subject: that so many philosophers ponder 'authority' strongly suggests that the Authority Thesis captures a source of puzzlement resembling that created by statements (1)–(3) above. It would not, otherwise, be understandable why philosophers, of all people, feel

themselves forced to explain what they call ‘authority.’ Why should philosophers care about ‘authority’ if it led to questions answerable by empirical means alone? The upshot is that no specification of the Authority Thesis can be considered adequate without, at least, the potential to create the kind of puzzlement that summons philosophers.

Because it will play a vital role in subsequent sections, allow me to highlight this point from a slightly different perspective. Philosophers have constructed subtle theories to explain what they call “authority.” For example, consider expressivism: the claim that self-ascriptions, rather than being public manifestations of introspective beliefs arrived at via a peculiar epistemic method, give direct expression to the mental states that they seem to concern. Sometimes, this position is amplified as the idea that the utterance ‘I am in pain’ is just a verbal substitute for crying and, thus, cannot be considered a truth-evaluable assertion. However, one need not endorse this rather bold claim to be an expressivist. Recently, for example, Dorit Bar-On—after distinguishing between at least two senses of expression, namely expression in the action sense (“expression<sub>1</sub>”) and expression in the semantic sense (“expression<sub>3</sub>”)—has argued that a “self-ascription expresses<sub>3</sub> the presence of [the] very same state that the avower expresses<sub>1</sub> in issuing the self-ascription” (Bar-On 2004: 265).<sup>7</sup> This claim preserves the notion that “avowals, unlike natural expressions, share logical and semantic properties with assertions, statements, reports, and the like” (Bar-On 2004: 245).

Of course, expressivism is not the only example of a philosophical theory intended to explain the special status of self-ascriptions. Another example is Donald Davidson’s account of what he calls “first-person authority,” which might be regarded as a transcendental argument along the following lines: we effectively communicate with each other; we cannot effectively communicate with one another unless the interpreter presumes that the speaker knows what their utterances mean; therefore, a

necessary condition for effective communication is that the interpreter presumes the speaker knows what their utterances mean. Further, Davidson argues, because knowing what one's utterances mean amounts to knowing what one thinks, a necessary condition for effective communication is that the interpreter presumes the speaker knows what they think and, thus, takes the speaker's self-ascriptions at face value.<sup>8</sup>

Undoubtedly, Bar-On's and Davidson's accounts do not come close to covering the multitudinous explanations that philosophers have provided. However, they are representative of the subtlety that these explanations typically exhibit. Now, the fact that philosophers respond to what they call 'authority' with fine-grained conceptual distinctions such as "expression<sub>1</sub>" and "expression<sub>3</sub>," or even with sophisticated transcendental arguments, strongly suggests that, initially, they were deeply puzzled by the observation to which the Authority Thesis gives expression. Philosophers seemingly perceive that this observation cannot be accounted for by empirical means alone, but deserves, instead, a kind of elucidation that only philosophical research can supply; this, I think, must somehow be reflected in any satisfactory specification of the Authority Thesis. Thus, we arrive at the third requirement again.

In the remaining sections of the paper, I will outline several initially promising specifications of the Authority Thesis and assess them in light of the three requirements formulated above. I will begin with what may be called the 'indubitability' interpretation of the Authority Thesis.

### **III. INDUBITABILITY**

On the indubitability interpretation, the Authority Thesis expresses the idea that self-ascriptions, as Bar-On writes, "are not easily subjected to doubt" (2004: 3).<sup>9</sup> Let us start the discussion with the following claim:



Indubitability<sub>1</sub> Necessarily, if a person *S* tells<sup>10</sup> another person *H* at time *t* that she (*S*) is currently in mental state  $\varphi$ , then *H* cannot rationally doubt at *t* that *S* is currently in  $\varphi$ —provided that *H* assumes that (i) *S* is sincere, (ii) *S* made no slip of the tongue, and (iii) *S* is not conceptually confused.

In my view, the most salient problem for Indubitability<sub>1</sub> is that we often suffer self-deceptions regarding some of our own dispositional mental states; compounding this difficulty, the fact we suffer from self-deceptions will sometimes be obvious to every person who knows us well except for ourselves.<sup>11</sup> A natural way to avoid this problem is to restrict the range of ' $\varphi$ ' to *conscious experiences*. However, even this modification will not protect Indubitability<sub>1</sub> from objections. Consider, for instance, the notorious example of the initiation trick in which a person is told she will be branded with a red-hot poker. However, instead of a red-hot poker, an ice cube is pressed against her bare stomach (Locke 1967: 86). To see why this example undermines the experience-restricted version of Indubitability<sub>1</sub>, suppose that, when the ice cube touches her stomach, the victim shouts, "I am in pain now!" Even if the people who press the ice cube to her skin assume she is sincere, made no slip of the tongue, and is not conceptually confused, they can rationally doubt that what she tells them is true—for it is reasonable to suppose that the victim is not in pain at the moment she shouts and that she, instead, merely senses the cold of the ice cube. Note that the issue is not whether the victim *actually feels* pain but rather whether her tormentors *have good reason to doubt* she is in pain at the moment they press the ice cube against her bare stomach. In my view, even someone who argues that the victim does feel pain must concede that the tormentors have good reason to doubt this.

The advocate of Indubitability<sub>1</sub> may react by invoking normality conditions. It seems natural, for example, to stipulate that the hearer assumes the self-ascription in question is made under *normal circumstances*. For instance, the speaker is assumed to be neither manipulated nor stirred up—let alone tortured—in any sense. However, there may also be reason to doubt Indubitability<sub>1</sub> under the assumption of perfectly normal circumstances. Consider the following:

The visual field is colored, at least for those of us who are not color-blind. It seems as if it is colored from one periphery to the other. Yet if you stare straight ahead while someone moves a colored playing card from directly behind you into your peripheral field, you are unable to say what color it is. Try it! The card can be moved surprisingly close to the center of your visual field before you can name the color. Since most of us know already that only the cones of the macula are sensitive to color while the rods on the rest of the retina are not, we should not be surprised by this experience. Yet we are surprised! It seems to us that the whole visual field is colored, though only the center of it really is. (Thompson 2009: 26)<sup>12</sup>

Now, imagine you ask a friend, who does not already know that only the center of our visual field is colored, the color that she currently experiences at the periphery of her visual field. Suppose that the answer is as follows: “I am experiencing a bright red there right now.” In my view, you have good reason to think that this statement is wrong. You are, at least *prima facie*, justified in believing that the periphery of your friend’s visual field is *not* colored. Of course, your friend may, surprisingly, have a different physiology from most of us and, in turn, actually have a visual field that is colored throughout. However, this mere possibility does not undermine your initial justification for doubting your friend’s self-ascription.

In my view, the best strategy for protecting the idea of indubitability from counterexamples is to restrict the range of self-ascriptions for which indubitability is

claimed: simply determine that only self-ascriptions to the effect that *one currently thinks that p*—given sincerity, etc.—are indubitable.<sup>13</sup> To recognize the point of this restriction, note that the verb ‘to think’ is used in a quite technical sense here. ‘*S* thinks that *p*’ does not imply that *S* is inclined to accept—let alone that *S* actually accepts—the proposition that *p*. Nor does it imply that *S* considers whether *p* is true. Rather, ‘*S* thinks that *p*’ simply means that *S* entertains—in whatever mode—the proposition that *p*.<sup>14</sup> Entertaining the proposition that *p* is an operation fundamental to all mental states whose content incorporates the proposition that *p*. It is not only present when a person believes that *p* but also when she, for example, wonders whether Ralph fears other people will think he believes it is not the case that *p*. Given this account of ‘to think that *p*,’ the following version of the indubitability thesis might be suggested:

Indubitability<sub>2</sub>    Necessarily, if a person *S* tells another person *H* at time *t* that she is currently thinking that *p*, then *H* cannot rationally doubt at *t* that *S* is currently thinking that *p*—provided that *H* assumes that (i) *S* is sincere, (ii) *S* made no slip of the tongue, and (iii) *S* is not conceptually confused.

I cannot find any counterexample. If the hearer *H* assumes that the speaker *S* is sincere, made no slip of the tongue, and is not conceptually confused, then *H* is rationally committed to the claim that *S* believes what she says, that is, that she believes herself to think that *p*. Now, recall that, given my account of ‘to think that *p*,’ *S* cannot believe that she thinks that *p* unless she thinks that *p*. In short, once *H* assumes that *S* is sincere, made no slip of the tongue, and is not conceptually confused, *H* is *rationally committed* to the claim that *S* thinks that *p*. This is why *H* cannot rationally doubt what *S* says. So, beyond doubt, Indubitability<sub>2</sub> is true. However, this is not to say that it satisfies the first requirement for any adequate specification of the Authority Thesis. Recall the first

requirement for an adequate specification: it must delineate the alleged authority as a feature that *at least a sufficiently large number of* self-ascriptions have. Now, self-ascriptions of the type ‘I am thinking now that  $p$ ’ form only a tiny fraction of self-ascriptions. Thus, Indubitability<sub>2</sub>, though true, does not seem to capture the phenomenon that philosophers have in mind when they speak of ‘authority.’

#### IV. PRESUMPTION OF TRUTH—FIRST ATTEMPT

In this section, I will discuss a specification of the Authority Thesis that seems more promising than the claim of self-ascriptions being immune to rational doubt. The idea is that self-ascriptions merely enjoy a *presumption of truth*—a presumption that does not attach to other-ascriptions. This idea is somewhat obscure, but I take it that an utterance by speaker  $S$  enjoys a presumption of truth if, and only if, the hearer  $H$  is justified in believing what  $S$  tells him, provided he lacks any good reason to doubt what  $S$  says.<sup>15</sup> As such, for  $H$  to be justified in accepting a given self-ascription, lacking good reason to doubt is already sufficient—positive evidence is not needed. Let us formulate this claim as follows:

Presumption of Truth<sub>1</sub>    If (1) a person  $S$  tells another person  $H$  at time  $t$  that she ( $S$ ) is currently in mental state  $\varphi$  and (2)  $H$  lacks any good reason to doubt what  $S$  says, then  $H$  is justified at  $t$  in believing that  $S$  is currently in  $\varphi$ —even if  $H$  possesses no positive evidence that  $S$  is currently in  $\varphi$  (aside from the fact that  $S$  tells him so).<sup>16</sup>

Presumption of Truth<sub>1</sub> seems true to me. Nonetheless, a difficulty arises. Recall that the Authority Thesis comprises two parts, holding that: (1) self-ascriptions enjoy a certain kind of authority, (2) which is *missing from other-ascriptions*. Therefore, to capture the

Authority Thesis completely, Presumption of Truth<sub>1</sub> must be flanked by a corresponding claim concerning other-ascriptions—the “complementary claim” (CC) to Presumption of Truth<sub>1</sub>:

CC to Presumption of Truth<sub>1</sub>      By contrast, if (1) a person *P* tells another person *H* at time *t* that a third person *S* is currently in mental state  $\varphi$  and (2) *H* possesses no positive evidence that *S* is currently in  $\varphi$  (aside from the fact that *P* tells him so), then *H* is not justified at *t* in believing that *S* is currently in  $\varphi$ —even if *H* lacks any good reason to doubt what *P* says.

Seemingly, then, the mere absence of any good reason to doubt might be sufficient for being justified in accepting a *self-ascription*, but *not* sufficient for being justified in accepting the corresponding *other-ascription*. In the latter case, *H* must not only lack any good reason to doubt but also have some evidence that *S* is currently in a particular mental state—the mere fact that *P* tells him so is not enough.

In my view, the CC to Presumption of Truth<sub>1</sub> is false. To be justified in accepting an other-ascription, the hearer does not need positive evidence. As in the case of self-ascriptions, it is sufficient that the hearer lacks any good reason to doubt. To substantiate this point, consider the debate surrounding testimonial justification. This debate stems from a fact that seems puzzling: we permanently accept a variety of assertions made by our fellow human beings, despite seemingly lacking any evidence to confirm what they tell us. Take the simple case of a tourist asking a stranger the way to the train station. If the stranger tells the tourist to turn left at the next corner, the tourist trusts her, even though he does not have the slightest idea of where the station is or whether the stranger is trustworthy. What is philosophically interesting, here, is the strong intuition to the effect that the tourist is *epistemically justified* in believing what

the stranger tells him about the location of the station. However, measured against traditional standards of justification, this intuition seems unfounded: according to traditional standards, being justified in believing that  $p$  requires positive evidence for  $p$ . Much of the debate surrounding testimonial justification thus turns on the question of *why* we are justified in accepting a stranger's say-so in situations similar to the tourist case.

Let's return to the question of whether the CC to Presumption of Truth<sub>1</sub> is true. To see that it is not true, simply consider the following situation. Suppose that a philosophy department that I have never visited invites me to a talk. Although I know the colleagues in this department by name, I do not know them personally, except for one whose name is N.N. However, I have not seen N.N. in a long time (and thus lack knowledge of his wellbeing). Regardless, as I enter the lecture room, I notice that N.N. is not there. Thus, I ask the first available person whether she knows N.N.'s whereabouts. To my horror, she replies, "Oh, didn't you know? N.N. collapsed in his room and is lying there, suffering a stabbing pain in his chest. The ambulance will arrive any moment." Let us suppose that I lack any good reason to doubt what the speaker says: first, the scenario that the speaker depicts is not unlikely—heart attacks can happen to anyone at any time; second, there is no indication that the speaker is insincere—she speaks in a tone and with a facial expression that is appropriate in this situation; third, I do not possess any evidence to the effect that N.N. is *not* in pain now—such as I would, for example, if I saw him cheerfully waving his hand from the opposite side of the lecture hall. However, I also lack any positive evidence. For example, I do not see him lying on the floor of his room, groaning and pressing his hands to his chest. Furthermore, I cannot rule out that the speaker is playing a joke on

me or that she is mentally ill and—no matter what question is directed to her—*always* says that N.N. is in pain.

The crucial point is as follows: if it is generally acknowledged that *the tourist* is justified in accepting what the stranger says, then it cannot be denied that *I*, too, am justified in accepting what the first available person says—namely, that N.N. is in pain right now. If it is acknowledged that I am justified in this belief, then it must be conceded that the CC to Presumption of Truth<sub>1</sub> is false—for we have a case in which a person is epistemically justified in believing an other-ascription despite lacking any positive evidence. If the CC to Presumption of Truth<sub>1</sub> is false, then the second requirement for any satisfactory specification of the Authority Thesis—which demands that other-ascriptions lack authority—is violated. One might say it is true that self-ascriptions enjoy a presumption of truth—but other-ascriptions enjoy it as well.

One might raise the objection that my argument presupposes a highly controversial view about the nature of testimonial justification, namely, that positive evidence is not necessary for testimonial justification: the *liberal view*. By contrast, proponents of what may be called the *conservative view*<sup>17</sup> claim that positive evidence *is* necessary for testimonial justification. Thus, conservatives typically argue that, despite appearances, the tourist who asks a stranger the way to the train station indeed possesses positive evidence. The tourist may have, for example, “a substantial amount of inductive evidence for believing that people are generally both sincere and competent when providing directions in normal contexts” (Lackey 2008: 181). If one analyzes the tourist case in this way, one is committed to analyze the N.N. case similarly. One may say, for instance, that in the N.N. case, I have a substantial amount of inductive evidence for believing that academics are generally both sincere and competent when providing information about the health of their colleagues in normal contexts. Thus

framed, the N.N. case does not raise any challenge to the CC to Presumption of Truth<sub>1</sub>—for the N.N. case cannot qualify as a situation in which a person is epistemically justified in believing an other-ascription despite lacking any positive evidence. Hence, the CC to Presumption of Truth<sub>1</sub> can be salvaged. Thus, the objection to my argument against the CC to Presumption of Truth<sub>1</sub> amounts to the following: “Whether the CC to Presumption of Truth<sub>1</sub> is false depends on whether liberalism concerning testimonial justification is true. However, you did not make any effort to establish liberalism; instead, you simply *presupposed* it. As such, your argument is deficient at the very least.”

The objection is at least partly right: the alleged falseness of the CC to Presumption of Truth<sub>1</sub> depends on the truth of liberalism, which I simply presupposed. However, this does not mean that the specification of the Authority Thesis at hand can be salvaged. Assume that the conservative view about testimonial justification is true, implying that the mere absence of any good reason to doubt is *not sufficient* for being justified in accepting a piece of testimony: the hearer must also have evidence that supports what the speaker says. Now, note that self-ascriptions are, themselves, pieces of testimony: one may contend that, similar to the stranger who informs the tourist about the way to the train station, a speaker who says “I am currently in pain” informs the hearer about her current mental state.<sup>18</sup> Thus, proponents of the conservative view are committed to the claim that a hearer without a single piece of confirming evidence is *not* justified in accepting a given self-ascription—even if he lacks any good reason to doubt. Consequently, proponents of conservatism are committed to denial of Presumption of Truth<sub>1</sub>.

To substantiate this point, consider the manner in which proponents of conservatism typically argue that positive evidence is necessary for testimonial



justification. The standard strategy involves considering situations in which all the requirements for testimonial justification generally accepted by liberals are satisfied but in which the hearer, nonetheless, lacks testimonial justification. In this sense, Lackey (2008: 168) imagines a case involving Sam, an average human being, who picks up a diary previously dropped by someone whom Sam correctly identifies as an alien. Because the first sentence of the diary reads, “Immediately after we landed on Earth, some of us were eaten by tigers,” and because, as Lackey stipulates, nothing about the diary gives Sam any good reason to doubt what its author writes, Sam eventually comes to believe that tigers have eaten some of the author’s fellow aliens. However, Lackey makes every effort to guarantee that Sam lacks any positive evidence to support the alien’s testimony:

[H]e has no common-sense alien-psychological theory, he has no beliefs about the general reliability of aliens as testifiers, he has no beliefs about the reliability of the author of this book, he has no beliefs about how ‘diaries’ function in this alien society, and so on. Moreover, if Sam attends to the narrative voice of the author in the hope of trying to assess her competence and sincerity, he would be engaged in a fruitless activity since there is no reason to believe that signs of competence and sincerity on the planet in question correspond to these signs on Earth. Sam cannot even compare the content of the reports in this diary to his background beliefs since he does not know that the words in this book are used in the same way that we on Earth use them. (Lackey 2008: 169)

The crucial question, asserts Lackey, is whether Sam is *justified* in believing that tigers have eaten some of the author’s fellow aliens. Lackey’s answer is negative: “[I]t seems plainly irrational for [Sam] to form the belief in question ... it seems clear that the appropriate epistemic response is to withhold belief” (Lackey 2008: 169 f.).

Regardless of one’s views on Lackey’s answer, it should be obvious that she would be committed to exactly the same answer if the scenario were slightly modified.

Suppose that the first sentence of the alien's diary read, "*I really feel sad because some of us were eaten by tigers immediately after we landed on Earth.*" According to her own standards, Lackey must admit it is irrational for Sam to believe that the diary's author feels sad because Sam still has no positive evidence to support the alien's testimony: he has no common-sense alien-psychological theory; he has no beliefs about aliens' general reliability as testifiers; he has no beliefs about the reliability of the diary's author; and so on. Therefore, I conclude that Lackey and other proponents of conservatism cannot accept Presumption of Truth<sub>1</sub>.

The ultimate outcome, here, is that claiming the Authority Thesis is best captured by Presumption of Truth<sub>1</sub> and its CC faces the following dilemma: either positive evidence is necessary for testimonial justification or it is not. If positive evidence is necessary, then Presumption of Truth<sub>1</sub> proves false, although its CC might be correct; if positive evidence is not necessary, then Presumption of Truth<sub>1</sub> might be true, but its CC proves false. Whichever way it is perceived—whether from a conservative or a liberal standpoint—the current proposal does not satisfy all the requirements for an adequate specification of the Authority Thesis: it violates either the first or the second requirement.

## V. PRESUMPTION OF TRUTH—SECOND ATTEMPT

In view of the difficulties presented in the preceding section, it might be asserted I have misconstrued the idea of the presumption of truth. It is not that the absence of any good reason to doubt is already sufficient for the hearer to be *justified* in accepting a self-ascription; rather, the idea is that, lacking any good reason to doubt and any positive evidence, the hearer's willingness to question a self-ascription is typically

lower than his willingness to question an other-ascription.<sup>19</sup> This idea might be expressed by the following dual claim:

Presumption of Truth<sub>2</sub>                      If (1) a person *S* tells another person *H* at time *t* that she (*S*) is currently in mental state  $\varphi$  and (2) *H* lacks any good reason to doubt what *S* says, then, given normal circumstances, *H*'s willingness to question whether *S* really is in  $\varphi$  is rather low—even if *H* possesses no positive evidence that *S* is currently in  $\varphi$  (aside from the fact that *S* tells him so).

CC to Presumption of Truth<sub>2</sub>                      By contrast, if (1) a person *P* tells another person *H* at time *t* that a third person *S* is currently in mental state  $\varphi$  and (2) *H* possesses no positive evidence that *S* is currently in  $\varphi$  (aside from the fact that *P* tells him so), then, given normal circumstances, *H*'s willingness to question whether *S* really is in  $\varphi$  is not as low—even if *H* lacks any good reason to doubt what *P* says.

Regarding willingness to question an ascription in the absence of both positive evidence and good reason to doubt, whether this is lower for a self-ascription than an other-ascription is a psychological question that requires empirical research. I am unaware of any empirical study that answers this question. So, Presumption of Truth<sub>2</sub> (including its CC) is empirically uncertain at least. However, even if Presumption of Truth<sub>2</sub> (including its CC) were true, it could not constitute an adequate specification of the Authority Thesis because it violates the third requirement, according to which any satisfactory specification must delineate the alleged authority as a feature that gives rise to a kind of puzzlement distinctive of philosophical problems.

To appreciate this point, ask yourself whether it is possible to create an apory from Presumption of Truth<sub>2</sub> and its CC that has the same force as the apory by which the mind-body problem is defined. Try this:

(A) In the absence of both evidence and good reason to doubt, there are no differences in our willingness to question statements similar in content but that are uttered by different persons.

(B) Self-ascriptions and corresponding other-ascriptions are similar in content but uttered by different persons.

(C) Our willingness to question a self-ascription in the absence of both positive evidence and good reason to doubt is lower than our willingness to question a corresponding other-ascription in the same circumstances.

Formally seen, this is an apory: the truth of any two of these statements implies the falsity of the remaining third. Nonetheless, the apory at hand does not have remotely the same force as the apory by which the mind-body problem is defined. The reason is that (A) is simply counterintuitive: why should we believe that, in the absence of both evidence and good reason to doubt, there are no differences in our willingness to question statements similar in content but uttered by different persons? We all know that our willingness to question a given statement is strongly influenced by temperament, prejudice, bias, social pressure, and other factors that are often beyond rational control. Thus, that the extent of our willingness to question a statement might depend on *who* utters it should not astonish us: it is perfectly compatible with existing knowledge about human psychology. Since there is no apory, there is no reason to consult philosophers.<sup>20</sup>

There is another route to essentially the same conclusion. Presumption of Truth<sub>2</sub> describes a fact concerning a peculiar type of behavioral disposition that people display in certain circumstances: their willingness to let self-ascriptions go

unquestioned is, *ceteris paribus*, higher than their willingness to let other-ascriptions go unquestioned. If this were the fact that philosophers typically attempt to explain, then their theories should address the following question:

(Q) Why is people's willingness to let self-ascriptions go unquestioned in the absence of both positive evidence and good reason to doubt higher than their willingness to let other-ascriptions go unquestioned in the same circumstances?

To avoid possible misunderstandings, note that (Q) should be understood as asking for the reasons that *actually motivate* the agent, i.e., reasons that are *psychologically effective* in the agent. Note, moreover, that motivating reasons in this sense should not be confused with justifying reasons, i.e., reasons that justify an action, as judged by a well-informed, impartial observer, whether they are psychologically effective or not. Now, philosophical explanations of the Authority Thesis typically provide justifying, not motivating, reasons. Take, for example, Davidson's account of first-person authority. According to Davidson, the presumption that the speaker knows what she means by her utterances—and thus knows what she thinks—is a necessary condition for effective communication. So, at least from the perspective of a well-informed, impartial observer, agents are *pragmatically justified* in letting self-ascriptions go: were they not to let them go unquestioned, they could not effectively communicate with each other. However, it is doubtful whether this linkage between effective communication and the practice of letting self-ascriptions go unquestioned is the psychologically effective reason in agents who let self-ascriptions go unquestioned. Most people simply do not know, let alone believe, that letting self-ascriptions go unquestioned is a necessary condition for effective communication. How could something of which the agent is totally unaware cause them to take a certain course of action? Thus, I consider it

uncharitable to interpret Davidson as providing a *motivating* reason for our practice of letting self-ascriptions go unquestioned. It is more charitable to interpret him as providing a *justifying reason* for that practice. If this is true, then Davidson's theory does not respond to (Q); rather, it responds to *another* question:

(Q\*) Why is it *appropriate* for people to let self-ascriptions go unquestioned in the absence of both positive evidence and good reason to doubt, whereas it is *not appropriate* to let other-ascriptions go unquestioned in the same circumstances?<sup>21</sup>

(Q\*), however, does not correspond to Presumption of Truth<sub>2</sub> (and its CC)—for Presumption of Truth<sub>2</sub> does not contain any *normative* element: it does not posit as *appropriate* people letting self-ascriptions go unquestioned in the absence of both positive evidence and good reason to doubt (though, certainly, it neither *denies* it is appropriate to let self-ascriptions go unquestioned). It simply states that, *as a matter of fact*, people's willingness to let self-ascriptions go unquestioned in the absence of both positive evidence and good reason to doubt is comparatively high. Thus, interpreting the Authority Thesis along the lines of Presumption of Truth<sub>2</sub> (and its CC) cannot help us understand why philosophers respond to what they call 'authority' in the way they actually do; in particular, why they neglect answering (Q) but instead seek answers to (Q\*). It is natural to suppose this is because (Q), unlike (Q\*), is of minor philosophical interest. As earlier acknowledged, that our willingness to question a given statement varies from case to case is perfectly compatible with existing knowledge about our psychology. Even without yet knowing exactly which factors account for that fact, we can be sure there is *some* empirical explanation, whether it will be provided by biology, psychology, sociology, or some other empirical science. In short, Presumption of Truth<sub>2</sub> does not give rise to the kind of puzzlement distinctive of philosophical problems.

Consequently, Presumption of Truth<sub>2</sub> does not satisfy the third requirement for an adequate specification of the Authority Thesis.

It might be argued that we can put things straight by modelling a specification of the Authority Thesis that corresponds to (Q\*). However, this will not succeed. Recall that we have already discussed promising specifications that correspond to (Q\*), namely Indubitability<sub>1</sub> and Presumption of Truth<sub>1</sub>, and found them inadequate: they violate the first or the second requirement for any adequate specification of the Authority Thesis. Thus, it seems that anyone seeking an adequate specification of the Authority Thesis faces a dilemma: either one arrives at a specification that is philosophically interesting but false (Indubitability<sub>1</sub>, Presumption of Truth<sub>1</sub> including its CC), or one arrives at a specification that is true but of minor philosophical interest (Indubitability<sub>2</sub>, Presumption of Truth<sub>2</sub> including its CC).

## VI. CONCLUSION

Let us reflect on the arguments thus far. According to many philosophers, the “cardinal problem of self-knowledge” (Wright 1998: 106) consists in explaining why self-ascriptions enjoy a kind of authority that other-ascriptions lack, that is, explaining why the Authority Thesis holds. This paper’s aim was not to provide an explanation but to ask whether there is a satisfactory specification of the Authority Thesis. For this purpose, I formulated three requirements that any satisfactory specification must meet. First, the specification must be such that the alleged authority is something that a sufficiently large number of self-ascriptions actually possess. Second, the specification must be such that the corresponding other-ascriptions lack this alleged authority. Third, the specification must be such that it gives rise to the kind of puzzlement distinctive of philosophical problems. I then discussed four suggestions for how the Authority Thesis

might be specified: Indubitability<sub>1</sub>, Indubitability<sub>2</sub>, Presumption of Truth<sub>1</sub>, and Presumption of Truth<sub>2</sub> (including their respective CCs). Indubitability<sub>1</sub> and Indubitability<sub>2</sub> were found to violate the first requirement, whereas Presumption of Truth<sub>2</sub> was found to violate the third requirement. The situation with Presumption of Truth<sub>1</sub> is more complicated as it leads to different difficulties depending upon whether a conservative or a liberal view is taken regarding testimonial justification. From a conservative perspective, Presumption of Truth<sub>1</sub> violates the first requirement, but, from a liberal perspective, Presumption of Truth<sub>1</sub> violates the second requirement.

Undoubtedly, these findings might not definitively establish that the Authority Thesis cannot be adequately specified. There might be other specifications that I did not consider. For example, some might object that I did not consider it might be specified as: *S*'s self-ascriptions are (or are treated as) more reliable, or more strongly justified, than anyone else's ascriptions of mental states to *S*.<sup>22</sup> However, I doubt that this specification is a good candidate for the Authority Thesis. First, consider the suggested specification *without* the parenthetical phrase 'or are treated as.' Because utterances are reliable or justified only in a derivative way—they are reliable or justified by virtue of the reliability or justification of the *beliefs* that they express—the specification amounts to a privileged access thesis along the following lines: '*S*'s beliefs about her own current mental states are more reliable, or more strongly justified, than anyone else's beliefs concerning *S*'s current mental states.' Now, recall that, earlier in this paper, I decided to take the Authority Thesis as one about pragmatic properties of utterances, not epistemic properties of beliefs. Thus, the claim at hand is not suitable as a specification of the Authority Thesis. Considered in isolation, it might be true and even philosophically interesting. However, it simply falls outside the scope of this paper.



Second, consider the suggested specification *with brackets removed*, namely: *S*'s self-ascriptions are treated as more reliable, or more strongly justified, than anyone else's ascriptions of mental states to *S*. This formulation is certainly suitable as a candidate specification of the Authority Thesis, as it describes a pragmatic property of utterances. However, it seems to suffer from the same defect as Presumption of Truth<sub>2</sub>—for it describes a behavioral pattern that is perfectly compatible with existing knowledge about human psychology. Thus, it does not yield the kind of puzzlement distinctive of philosophical problems. Consequently, it violates the third requirement of any adequate specification of the Authority Thesis.

Let me address a final concern. If the Authority Thesis cannot be adequately specified, it needs to be explained why so many philosophers have mistakenly considered this possible. The answer lies in an ambiguity error. Consider, for example, Davidson's starting point (1984: 101): "When a speaker avers that he has a belief, hope, desire or intention, there is a presumption that he is not mistaken, a presumption that does not attach to his ascriptions of similar mental states to others." This sentence is notoriously ambiguous. On one reading, it asserts that, *as a matter of fact*, self-ascriptions are treated by their addressees with a certain kind of deference that is not to be found in the case of other-ascriptions. On another reading, it declares that addressees are *entitled* to treat self-ascriptions with a certain kind of deference with which they are *not entitled* to treat other-ascriptions. On the first interpretation, Davidson's sentence is along the lines of Presumption of Truth<sub>2</sub> (including its CC), which, as we have already seen, may be true but is not especially problematic in philosophical respects. On the second interpretation, Davidson's sentence is along the lines of Presumption of Truth<sub>1</sub> (including its CC), which may generate a philosophical problem but is false. Now, I suppose that the widely held belief in an adequate specification of the Authority Thesis

is due to a confusion of these two interpretations. Regarding the first and the second requirement of any satisfactory specification of the Authority Thesis, Davidson's sentence, or some analogue,<sup>23</sup> is interpreted according to the *first* reading. However, regarding the third requirement, the sentence is interpreted according to the *second* reading. This oscillation between two interpretations gives the impression that all three requirements are satisfied by one and the same specification of the Authority Thesis. Actually, however, there are *different* specifications involved, each of which violates at least one requirement.

Even if this paper's considerations cannot force us to conclude there is no adequate specification of the Authority Thesis, at the very least, the existence of such a specification cannot be taken for granted. Thus, every philosopher who seeks to explain the phenomenon (allegedly) captured by the Authority Thesis should strive to specify in detail what this phenomenon is supposed to be—and why it needs to be addressed by *philosophers*. Hackneyed phrases such as 'self-ascriptions cannot easily be challenged' or 'usually, self-ascriptions are presumed to be true' are insufficient for this purpose. Thus, I hope that my arguments will contribute to shifting the burden of proof, moving the onus from those who *deny* to those who *maintain* there is an adequate specification of the Authority Thesis.<sup>24</sup>

## ENDNOTES

<sup>1</sup> Wright lists two other features of self-ascriptions that warrant explanation: "transparency" and "groundlessness."

<sup>2</sup> See Gertler 2011: 3.

<sup>3</sup> See Alston 1971 for a detailed taxonomy and systematization of privileged access theses.

<sup>4</sup> Thus, authority in the sense of the Authority Thesis amounts to what Fricker 2000 calls 'language-game authoritativeness'. See also Lawlor 2003: 549.

<sup>5</sup> LePore and Loewer 1987: 630.

<sup>6</sup> I am well aware that this view of the nature of philosophy is not universally shared; however, I believe it captures the essence of what many philosophers think about their discipline. Thus, I will not argue for the metaphilosophical view adopted here, but simply presuppose it. If one disagrees, one may take my line of

---

argument in this paper in a conditional sense: given the traditional metaphilosophical view, there is no satisfactory interpretation of the Authority Thesis.

<sup>7</sup> See also Tugendhat 1986, Lecture 6, for a similar claim.

<sup>8</sup> Cf. Davidson 1984. See Hacker 1997: 290, for a reconstruction of Davidson's argument along these "transcendental" lines.

<sup>9</sup> The view that self-ascriptions enjoy some kind of (conditional) indubitability, albeit not uncontroversial, is widely held in the literature. See Heal 2002.

<sup>10</sup> I take 'to tell' as an achievement verb in the sense it implies that the hearer *understands* what the speaker says.

<sup>11</sup> Compare the case of the sexist professor described in Schwitzgebel 2012: 191, to give just one of the many examples adduced in the literature.

<sup>12</sup> The example originally appeared in Dennett 1991: 53.

<sup>13</sup> An anonymous referee has suggested there might be a second option that seems promising here: restricting the range of self-ascriptions for which indubitability is claimed to those that express *direct phenomenal beliefs* in the sense of Chalmers 2003. Direct phenomenal beliefs are said to literally contain the conscious experiences they concern and, thus, seem to be infallible. Consequently, self-ascriptions that express direct phenomenal beliefs cannot rationally be doubted. However, the problem with this suggestion is that an utterance of 'I am experiencing a bright red there right now' does not reveal whether it expresses a direct phenomenal belief; therefore, there is adequate room for doubt in the hearer. Arguably, one might fix this problem by stipulating that the hearer, in addition to assuming the speaker is sincere, made no slip of the tongue, and is not conceptually confused, assumes that the speaker's utterance expresses a direct phenomenal belief. However, this assumption makes the suggestion rather trivial, as the infallibility of direct phenomenal beliefs "is an immediate consequence of [their] definition" (Chalmers 2003: 242).

<sup>14</sup> See Kemmerling 2012: 402.

<sup>15</sup> The idea of elucidating the notion of authority as a presumption of truth originated with Davidson 1984. However, Davidson does not clarify whether his claimed presumption of truth is merely a description of a psychological matter of fact or should be read in a normative sense, according to which the addressee of a given self-ascription is *entitled* to presume that the speaker is not mistaken. This ambiguity will play a role in this paper's later sections, but, for now, I assume the normative reading. In that regard, I am in accord with the majority of contemporary scholars. See, for example, McGeer 2007 and Medina 2006.

<sup>16</sup> By 'positive evidence,' I mean the set of propositions to which *H* has access—typically from observing the behavior of *S*—that support the proposition that *S* is currently in  $\varphi$ . As pointed out by an anonymous referee, it is hard to conceive of any conversational situation in which the hearer does not have access to *some* propositions (aside from the proposition that *S* says she is currently in  $\varphi$ ) that bear on the trustworthiness of what the speaker says. Moreover, it seems that the situations I discuss subsequently in this section are ones in which the hearer does, indeed, have access to such propositions. Thus, it might be better to formulate the "even if"-condition of Presumption of Truth<sub>1</sub> thus: "... even if *H* possesses no positive evidence that *S* is currently in  $\varphi$  (aside from the fact that *S* tells him so) *or does not make use of any such positive evidence to which he might potentially have access.*" However, for the sake of brevity, I preserve my original phrasing.

<sup>17</sup> Cf. Green 2008 and Graham 2006 for similar uses of 'liberal' and 'conservative.'

<sup>18</sup> This contention is not wholly uncontroversial. According to expressivists such as Malcolm 1954 and 1958, for example, a speaker who says, "I am currently in pain" does not *assert* anything, but only *evinces* their pain. From this perspective, self-ascriptions are not pieces of testimony. Thus, the contention that self-ascriptions are pieces of testimony should not be read as a general statement about self-ascriptions, regardless of one's specific theoretical perspective on them. Therefore, this should just be read as a contention that should be shared by anyone accepting Presumption of Truth<sub>1</sub> or its Complementary Claim. I am grateful to an anonymous referee for raising this point.

---

<sup>19</sup> See Bar-On and Long 2003: 181.

<sup>20</sup> As pointed out by an anonymous referee, there seems to be a striking disanalogy between the instances that falsify thesis (A) of the wannabe-apory and the situation that holds true according to Presumption of Truth<sub>2</sub> (including its CC). Regarding the topic of *S*'s current mental states, Presumption of Truth<sub>2</sub> (including its CC) states that *S* enjoys hearer's deference, while anyone other than *S*, speaking about the same topic, will not. There does not seem to be any analogous case in other areas. Of course, there might be topics on which certain *sorts* of people enjoy hearer's deference—as is the case with, for example, mothers who discuss the topic of motherhood. However, are there really any topics with regard to which one *and only one* person enjoys hearer's deference? Thus, it seems that we have, here, at least the seed for generating a new, and this time genuine, apory. However, I do not think that there is any grave disanalogy. We are familiar with situations of topics with regard to which just one person enjoys hearer's deference. Consider the relationship between religious leaders and their devotees, for example. However, even if there were a disanalogy, it would not be detrimental to the point I wish to make, namely that, first, any differences in our willingness to question equipollent statements uttered by different persons *must be traceable to some factor or other that is psychologically effective*, and, second, that this factor, qua its psychological effectiveness, *must be empirically ascertainable*. Therefore, there is simply no reason to turn to philosophy here, given that the task of philosophers is to approach problems irreconcilable by empirical means alone.

<sup>21</sup> Similar considerations apply to expressivism. According to Bar-On, “there is no more mediation between the ‘I’-ascriptions we issue when avowing and the conditions we self-ascribe than there is between non-self-ascriptive expressions and the conditions they express” (2004: 399). Again, this account might well explain why it is *appropriate* to let self-ascriptions go unquestioned. However, it does not determine the factors that are psychologically effective when we do so. Thus, like Davidson, Bar-On addresses (Q\*), not (Q): she provides a *justifying* reason for letting self-ascriptions go unquestioned, not a *motivating* one.

<sup>22</sup> I owe this objection to an anonymous referee. I thank them for pressing me on this point.

<sup>23</sup> Please note that I do not intend to criticize Davidson *in particular*. I cite Davidson because he provides an especially clear example of a sentence that is ambiguous between the two outlined readings. There are many formulations in the literature as ambiguous as Davidson's. Compare, for example, Bar-On's claim that self-ascriptions are rarely questioned or corrected or Wright's claim that self-ascriptions are accredited with authority: these claims are likewise amenable to both a descriptive and normative interpretation. Thus, all that I state here about Davidson is applicable to many other authors.

<sup>24</sup> This paper is based on a talk given by the author at an interdisciplinary workshop on self-ascriptions and self-knowledge, titled “think<sup>1</sup>,” at University College Freiburg, Germany in November 2014. I am indebted to the audience, especially to Wolfgang Carl, Simon Dierig, Christopher Gutland, Wolfgang Freitag, Andreas Kemmerling, Katharina Kraus, and Stephan Packard, for constructive criticism and valuable suggestions. Further, I wish to express my warmest thanks to Thomas Spitzley, who not only provided the freedom necessary for my research but also gave me the opportunity to present a previous draft to his graduate seminar at the University of Duisburg-Essen in January 2015. I am grateful to Anna-Maria Eder, Lars Dänzer, Philipp Kahlert, Insa Lawler, Thorsten Sander, Tobias Steinig, and Giuliano Torrenco for their valuable comments. Finally, I wish to thank the two anonymous referees of this journal for the goodwill with which they formulated their suggestions and criticism. The work on this article was supported by research grants BA 2269/2-1 and BA 2269/2-2 from the German Research Foundation (DFG).

## BIBLIOGRAPHY

- Alston, William. 1971. “Varieties of Privileged Access.” *American Philosophical Quarterly* 8: 223-241.

- Bar-On, Dorit. 2004. *Speaking My Mind: Expression and Self-Knowledge*. Oxford: Oxford University Press.
- Bar-On, Dorit and Douglas C. Long. 2003. "Knowing Selves: Expression, Truth and Knowledge." In *Privileged Access: Philosophical Accounts of Self-Knowledge*, ed. Brie Gertler. Aldershot: Ashgate, 179-212.
- Chalmers, David. 2003. "The Content and Epistemology of Phenomenal Belief." In *Consciousness: New Philosophical Perspectives*, ed. Quentin Smith and Aleksandar Jokic. Oxford: Clarendon, 220-272.
- Davidson, Donald. 1984. "First Person Authority." *Dialectica* 38: 101-112.
- Dennett, Daniel. 1991. *Consciousness Explained*. London: Penguin.
- Fricker, Elizabeth. 2000. "Self-Knowledge: Special Access versus Artefact of Grammar—A Dichotomy Rejected." In *Knowing Our Own Minds*, ed. Crispin Wright, Barry C. Smith, and Cynthia MacDonald. Oxford: Oxford University Press, 155-206.
- Gertler, Brie. 2011. *Self-Knowledge*. London: Routledge.
- Graham, Peter J. 2006. "Liberal Fundamentalism and its Rivals." In *The Epistemology of Testimony*, ed. Jennifer Lackey and Ernest Sosa. Oxford: Oxford University Press, 93-115.
- Green, Christopher R. 2008. "The Epistemology of Testimony." In *The Internet Encyclopedia of Philosophy*, accessed March 8, 2017, <http://www.iep.utm.edu/ep-testi/>.
- Hacker, Peter M. S. 1997. "Davidson on First Person Authority." *The Philosophical Quarterly* 47: 285-304.
- Heal, Jane. 2002. "On First Person Authority." *Proceedings of the Aristotelian Society* 102: 1-19.

- Kemmerling, Andreas. 2012. "First Person Authority without Glamorous Self-Knowledge." In *Epistemology: Contexts, Values, Disagreement*, ed. Christoph Jäger and Winfried Löffler. Frankfurt a. M.: Ontos, 401-433.
- Lackey, Jennifer. 2008. *Learning from Words: Testimony as a Source of Knowledge*. Oxford: Oxford University Press.
- Lawlor, Krista. 2003. "Elusive Reasons: A Problem for First-Person Authority." *Philosophical Psychology* 16: 549-564.
- LePore, Ernest, and Barry M. Loewer. 1987. "Mind Matters." *The Journal of Philosophy* 84: 630-642.
- Locke, Don. 1967. *Perception and Our Knowledge of the External World*. London: Allen & Unwin.
- Malcolm, Norman. 1954. "Wittgenstein's Philosophical Investigations." *The Philosophical Review* 63: 530-559.
- . 1958. "Knowledge of Other Minds." *The Journal of Philosophy* 55: 969-978.
- McGeer, Victoria. 2007. "The Moral Development of First-Person Authority." *European Journal of Philosophy* 16: 81-108.
- Medina, José. 2006. "What's so Special about Self-Knowledge?" *Philosophical Studies* 129: 575-603.
- Schwitzgebel, Eric. 2012. "Self-Ignorance." In *Consciousness and the Self: New Essays*, ed. Jeeloo Liu and John Perry. Cambridge: Cambridge University Press, 184-197.
- Thompson, David L. 2009. *Daniel Dennett*. New York: Continuum Books.
- Tugendhat, Ernst. 1986. *Self-Consciousness and Self-Determination*. Cambridge, MA.: MIT Press.

Wright, Crispin. 1998. "Self-Knowledge: The Wittgensteinian Legacy." *Royal Institute of Philosophy Supplement* 43: 101-122.