

Infallibility about the Self

GLEN HOFFMANN

Department of Philosophy. Faculty of Liberal Arts and Professional Studies York University
416 Ross South. Canada
ghoffman@yorku.ca

RECEIVED: 08-02-2013 • ACCEPTED: 16-05-2013

ABSTRACT: On a common formulation, rationalist infallibilism is committed to two main theses: (i) ‘analytic a priori infallibilism’ and (ii) ‘synthetic a priori infallibilism’. According to thesis (i), a relatively wide range of analytic a priori propositions can be infallibly justified. According to thesis (ii), a relatively wide range of synthetic a priori propositions can be infallibly justified. In this paper, I focus on rationalist infallibilism’s second main thesis, what is being called ‘synthetic a priori infallibilism’. I argue that synthetic a priori infallibilism, and by extension rationalist infallibilism, is untenable. In particular, exploring what seems to be the only potentially plausible species of synthetic a priori infallibility, I reject the infallible justification of propositions about the self.

KEYWORDS: Cogito – dubito – infallibility – privileged access – rationalism – self-knowledge.

1. Introduction: rationalist infallibilism and propositions about the self

On a traditional formulation, rationalist infallibilism is a doctrine committed to two main theses: (i) ‘analytic a priori infallibilism’ and (ii) ‘synthetic a priori infallibilism’. According to thesis (i), a relatively wide range of analytic a priori propositions can be infallibly justified. According to thesis (ii), a relatively wide range of synthetic a priori propositions can be infallibly justified. In sum, then, for the rationalist infallibilist two main categories of propositions are susceptible of infallible justification: (i) analytic propositions such as logical, mathematical, and ‘conceptual’ propositions, and (ii), synthetic propositions such as those concerning the external world

and the self. Though rationalist infallibilism is increasingly falling out of favor, there is no shortage of notable 20th-21st century adherents of one or more of the two species of infallible a priori justification. Among others, Fumerton (2001), McGrew (2003), and McGrew et al. (2007) still profess the a priori infallibility of some analytic propositions, while Burge (1986; 1988; 1996) and Lewis (1996) have recently defended the a priori infallibility of some synthetic propositions (see Hoffmann 2011, 241).

In this paper, I examine rationalist infallibilism's second thesis, so-called 'synthetic a priori infallibilism'.¹ More specifically, focus will be restricted to what are likely the only potentially viable candidates for synthetic a priori infallible justification, propositions about the self. In the course of this exploration, I reject the possibility of the infallible a priori justification of propositions about the self. Moreover, on the seemingly reasonable assumption that propositional knowledge involves propositional justification, that for a proposition to be known it must in some sense be justified, the implication is that infallible a priori *knowledge* about the self is also out of reach.² On the other hand, this anti-infallibilist argument does not eliminate the possibility of a priori justification and or knowledge (about the self or otherwise) *in general*, since it leaves open that some species of *fallible* a priori justification/knowledge might be within reach.

2. Infallibility about the self

One primary candidate for a priori infallibility, perhaps the most promising one, is in the domain of syntheticity.³ To name just a few, proposals

¹ I undertake a similar project in Hoffmann (2012).

² In what follows, though, discussion will primarily focus on *justification* instead of *knowledge*, in order to avoid questions surrounding the definition or analysis of the concept of *knowledge*.

³ I assume here that there is (i) an analytic/synthetic distinction and (ii) an a priori/a posteriori distinction that can be drawn along relatively traditional lines: (i) propositions whose truth depends exclusively on meaning/propositions whose truth doesn't depend exclusively on meaning, and (ii) propositions justifiable independently of experience/propositions not justifiable independently of experience. If it turns out (i) and/or (ii) is false, this is of little concern, since surely skepticism that there is a special domain of analytic and/or a priori propositions is far more likely to be harmful to infallibilist than anti-infallibilist aspirations.

for synthetic a priori infallible justification have included propositions about material objects, substances, properties, and the self (see Hoffmann (2011, 246). For example, the first three proposals for synthetic a priori infallible justification have been canvassed in some form or another by both Price (1953) and Lewis (1996). There seems to be a consensus, though, that these kinds of proposals for synthetic a priori infallibility have not survived sustained review. In any event, whatever the merits of the case for synthetic a priori infallibility about various external world phenomena, discussion here is confined to a class of synthetic propositions whose a priori infallibility still seems to be a live issue – propositions about the self.⁴

Traditionally, propositions about the self have been at the heart of the infallibilist program in epistemology. For the infallibilist, (some) propositions about the self are supposed to be self-justifying or self-verifying in the basic sense that their mere belief establishes their truth. In specifically rationalist terms, a proposition about the self is supposed to be self-justifying in the sense that its belief is *a priori* sufficient to establish its truth. Descartes' proposals about first-person mental states (1996/1641, 80ff) are the *locus classicus* for a priori infallibility about the self (see Hoffmann 2011, 47; 2012, 57f):

(Cogito) I am thinking (therefore I exist as a thinking thing)

and along the same lines

(Dubito) I am doubting (therefore I exist as a doubting thing).

Independent of whether (Cogito) and (Dubito) are self-justifying at all, i.e., whether belief in them establishes their truth (in some potentially *fallible* sense), is the question of whether they are susceptible of *infallible* justification, i.e., whether belief in them establishes their truth infallibly or indefeasibly. For the infallibilist the pertinent question is whether Cogito-like propositions, or any other proposition about the self of significant cognitive interest, can be justified in a way that is truth-entailing? More spe-

⁴ Other proposals for synthetic a priori infallibility include putative ontologically necessary truths to the effect that ⟨something exists⟩ or ⟨at least one thing exists⟩, and weakened Cogito-like propositions such as ⟨thought exists⟩ or ⟨thinking exists⟩. I leave open that one or more of these synthetic propositions is an infallibly justified a priori truth, while bearing in mind their indisputably negligible cognitive import in comparison with classical Cogito-like propositions (to be considered next in the text).

cifically, can these propositions be justified in a way that confers on them absolute warrant, or a probability of 1, where ‘probability’ is read in putatively *objective* terms, and ‘truth’ is read in standard realist terms?⁵

One rationale for rejecting infallible a priori self-justifying propositions stems from a firmly entrenched epistemological principle. The received view at present seems to be that justification for propositions about the self can never be *determinately* reflexive or self-referring (e.g., McDonald 1998). For example, if I have the thought ⟨I am currently thinking⟩, the warrant for the truth of the proposition expressed by this thought cannot be obtained by the mere having of this thought. Instead, it is reasoned, it must be secured via some kind of reflection on the thought. Moreover, insofar as reflection implies the existence of two, numerically distinct thoughts present in this case (a first and a second-order thought), it is unclear how the second-order thought ⟨I think that I am thinking⟩ can infallibly justify its first-order constituent ⟨I am thinking⟩. The upshot of this line of argument is that Cogito-like propositions such as ⟨I am thinking⟩, or even iterated Cogito-like propositions such as ⟨I think that I am thinking⟩, do not appear susceptible of infallible justification.

But this nascent anti-infallibilist line of argument is less than decisive as presently formulated since the Cartesian infallibilist about the self has at her disposal a steadfast reply to it. For Descartes (Cogito) and (Dubito) are infallible a priori self-justifying propositions since they are instances of *direct* ratiocination that supply unmediated acquaintance with the reality of the thinking self (Descartes 1996/1641, 80–81). David Lewis (1996, 564ff) similarly seems to construe (Cogito) and (Dubito) as pure rational intuitions, minus the Cartesian metaphysics (about subjects) and commitment to any form of internalism, intuitions that furnish unmediated access to the reality of the subject (see Hoffmann 2011, 248; and 2012, 57ff for more on this).⁶

⁵ More on this construal of *infallibility* and its bearing on the possibility of a priori infallibility about the self in §5.

⁶ One might object to direct, time-sensitive rational intuitions as a source of justification/knowledge on *pro tanto* empirical grounds. If one’s direct ‘acquaintance’ with subjective reality is restricted to the specious present and vanishes on second-order reflection, it might be argued, Cartesian rational intuitions cannot confer rational justification on propositions about first person mental states (or any other content-bearing state). Assuming judgment is diachronic and rational intuitions are synchronic, intuitions indexed to time fragments lack normative force since they disappear just as one registers

Needless to say, no claim to the unlimited scope of infallible a priori justification about the self is made here.⁷ For Descartes and Lewis the possibility of infallibility about the self is conditional upon the satisfaction of explicit desiderata, respectively, Descartes' clearness and distinctness criterion and Lewis' numerous criteria for ignoring counterfactual possibilities—states of affair that would undermine the truth of the proposition at issue. But in situations where the relevant criteria are met one can be acquainted with one's mental states directly and immediately,⁸ in a way that eliminates the need for second-order reflection. For Descartes, although Lewis is more circumspect on this point, this includes mental states with propositional content, such as beliefs about the self (e.g., ⟨I am now thinking⟩). It follows, on this general outlook, that there is at least a small class of propositions about the self that are determinately reflexive, and consequently, whose belief guarantees their truth (see also Burge 1986; 1988).

a judgment about them. I reserve judgment on this kind of objection since it enlists controversial premises concerning the nature of memorial cognition and knowledge, e.g., that they are temporal tokens, indexed to their time of conception. For discussion of the nature of memorial cognition and knowledge and its bearing on the viability of first person a priori judgments, see Boghossian (1989), Ludlow (1999), Brueckner (1997), and Goldberg (1997).

⁷ For ease of exposition, 'infallibility about the self' will hereafter largely be used as shorthand for 'infallible a priori justification about the self' and 'infallible a priori knowledge about the self'.

⁸ The general notion of direct a priori acquaintance demands close inspection. In particular, beyond indistinct formulations of the phenomenon in terms such as 'the light of reason', 'intuitive obviousness' or the like, important issues arise concerning the modal status, empirical status, and truth-conduciveness of pure rational intuitions. Typically, rational intuitions are deemed necessary, impervious to empirical data, and infallibly truth-conducive. For example, Bonjour (2005, 99), in deference to the first two criteria, declares rational intuitions provide "direct or immediate insight into the truth, indeed the necessary truth, of the relevant claim They are thus putative insights into the essential nature of things or situations of the relevant kind, into the way that reality in the respect in question *must* be". Turri (2011), contra the traditional view, defends the possibility of some kinds of *contingent* a priori rational intuitive knowledge. The bearing of such issues on the possibility of infallibility about the self will become especially evident in §4.

3. Infallibilism and direct ratiocination

In consequence, a more direct confrontation of Cartesian infallibilism about the self is required. As a first attempt, it might be argued against the Cartesian infallibilist that Cogito-like self-justifying propositions cannot be rationally justified *in esse*. At bottom, direct acquaintance with the mental is a notoriously perplexing epistemic phenomenon (see Bonjour 2003, 17f for a discussion of some of the main concerns): it is an enigma how brute mental states (about the self or otherwise) can be vehicles of rational justification. A pivotal concern is mental states of the form ⟨I'm experiencing like *this*⟩, where *this* designates a primitive feature of experience, look devoid of cognitive content—the kind of content that can figure in judgments.

To be sure, mental states of this kind lack *descriptive* content since they take the form ⟨I'm experiencing like *that*⟩ instead of ⟨I'm experiencing like *that* and not like *that*⟩. Such mental states, it can scarcely be denied, do not encode either definite or indefinite descriptions, on any remotely plausible construal of 'description'.⁹

The crucial question, then, is whether the mental states under consideration possess some kind of 'minimal', non-descriptive cognitive content, content that can play a justificatory role. One reason for being skeptical about this possibility is there doesn't seem to be an intelligible account of the reference of ostensibly primitive mental states. On the surface, the referent of the demonstrative *this* in such mental states is intractable. For starters, since it is fundamentally deictic, *this* obviously does not designate a context-independent state of affairs. More importantly, the referent of *this*, since it is necessarily a primitive feature of reality, looks to be completely *subject*-dependent, meaning it doesn't even seem to have a *context-relative* semantic value—a semantic value stable across some set of contextually relevant variables. If this is correct, the mental states in question can be viewed as purely phenomenal, and consequently, as lacking cognitive significance. It would follow, in this case, that they cannot furnish support for particular hypotheses or claims, that they cannot raise the credibility of

⁹ Some infallibilists (e.g., McGrew 2003; McGrew et al. 2007), typically of an internalist bent, explicitly reject this seemingly innocuous thesis. A more direct defense of this thesis is provided in §4, in our discussion of the infallibilist's appeal to the phenomenon of *direct reference* or *acquaintance* in application to beliefs about mental states.

a proposition, belief, or utterance in comparison with any other proposition, belief, or utterance.

Naturally, in response, the infallibilist will contend that the referent of the demonstrative *this* in primitive mental states is, in a fundamental sense, non-discursive: i.e., it lacks conceptual constituents, complex structure, discrete components, etc. In this connection, the infallibilist might appeal to a special kind of rudimentary cognition (the likes of which has been articulated by Husserl and Brentano, among others), a kind of cognition that is prior to language or possibly even to conceptualization, as the source of self-justifying infallible belief about the self. But the view that there can be pre-conceptual, or what has been called ‘semi-judgmental’ mental states that infallibly justify propositions about the self looks patently untenable. A pivotal concern is it seems virtually undeniable that mental states that are vehicles of justification must in *some* sense be representational (Bonjour defends this contention in Bonjour – Sosa 2003, 20). This, of course, is not to say that the representational content of justificatory mental states need be anything like a truth-evaluable propositional thesis or assertion, or that it be explicitly formulatable in terms of something like ‘comparison’ or ‘contrast’ classes. At a minimum, though, such mental states must have some kind of representational content (or perhaps *informational* content, if one prefers), i.e., they must depict things as being one way rather than another, in order to function as justifiers. But since the mental states in question lack such content, they cannot supply warrant for propositions about the self, let alone *absolute* warrant. The irresistible inference is that direct ratiocination about the self cannot be the spring of a priori infallibility about the self.¹⁰

4. Direct ratiocination and conceptual acquaintance

Another reply to our dismissal of infallibilism about the self involves the attempt to develop a more theoretically motivated account of direct ratioci-

¹⁰ The argument I advance here has some resemblance to a widespread anti-foundationalist argument according to which a mental state cannot both (i) have cognitive import and (ii) provide a foundation for propositional knowledge. Here I take no stance on the anti-foundationalist polemic. Since we haven’t ruled out the possibility that there are defeasible foundations for propositional knowledge that can be secured in a manner that does not involve direct acquaintance with mental states, our argument has no direct bearing on fallibilist variants of foundationalism.

nation. In brief, in this vein, one might defend the possibility of a kind of direct conceptual acquaintance on the basis of an account of first person rational intuitions, an account that might be considered in some sense phenomenologically faithful. For example, one might exploit a view of first person rational intuitions along the lines of Bonjour's (1996; 2003) and Bealer's (2000) *appearance* model to this effect, where such intuitions involve a kind of *seeing* or *visualizing* of subjective reality, rather than a conceptual or propositional grasp of it (n.b.: neither Bonjour nor Bealer are infallibilists about the self). Since on this model Cogito-like propositions take the general form $\langle I'm \text{ experiencing/feeling/being-appeared-to like } that, \text{ therefore } that \rangle$, where *that* does not refer to a propositional object, they might be thought to secure direct (non-discursive) acquaintance with the reality of the thinking/experiencing self. The infallibility of Cogito-like propositions in this case is thus supposed to stem from the fact that they enjoin a kind of direct reference or acquaintance with mental states (McGrew et al. 2007, 132; Burge 1986; Burge 1988), analogous to the way some have claimed one might be directly acquainted with features of the external environment through perception.

Importantly, this variant of infallibilism about the self is not tied to internalism (as in McGrew et al. 2007) – where the content of and/or the justification conditions for one's mental states are in some sense 'subject-internal' (e.g., introspectable, open to consciousness, transparent, luminous, etc.). A number of self-proclaimed externalists, including Burge (1986; 1988), Lewis (1996), Heil (1988), and Davidson (1984; 1987), have defended an account of a priori self-knowledge along the lines of the account currently being sketched (though seemingly only the former two can rightly be viewed as infallibilists). Burge (1988, 659-660), most notably, has developed an influential account of infallible a priori self-knowledge that is non-discursive in the sense under consideration, one in which there is a non-inferential link between conception and apperception of the self (and wherein the infallibility of certain beliefs (or thoughts) about the self is guaranteed by their reflexive character – by the fact that they 'logically lock' onto their first-order constituents).

In my view, though, there are powerful reasons for rejecting the refined formulation of direct mental acquaintance, whether internalist or externalist in orientation. But even if it is possible to directly apprehend the content of a mental state, or to have indefectible 'privileged access' to certain kinds of mental phenomena, this looks to be small consolation to the infal-

libilist. On our central line of argument, mental states with which one is directly acquainted *ipso facto* lack cognitive import since they are devoid of representational or informational content. Since mental states of the form ⟨it seems, appears, or feels to me that *such-and-such*⟩ concern “those conditions of the subject, whatever they are, which are accessible to the subject, whenever they obtain...” (Williamson 2000, 15), such states are veridical *by fiat*. Primitive mental states must, on this score, be considered strictly phenomenal, and quite possibly, beyond the limits of what is expressible.

Williamson (2000, 14ff), among others, has also argued for the epistemic triviality of primitive or what he calls ‘luminous’ mental states, on the grounds that they lack truth-tracking conditions ensuring their ‘safety’ or ‘secure’ application. Bypassing Williamson’s externalist-motivated *universal* anti-luminosity argument,¹¹ my principal contention is that primitive mental states, insofar as they concern the putative content of *immediate* experience or the character of *bare* sentiment, are cognitively vacuous. Whatever import they might have, such states are effectively nebulous from the standpoint of epistemology: they can neither figure in judgments nor confer rational justification on beliefs, statements, utterances, or any other bearer of propositional content.

Now obviously there are fundamentally divergent views about the nature, scope, and limits of cognitively significant mental content. While the argument currently being developed leaves open that there can be cognitively significant mental states (though not via a direct or unmediated grasp of their content), no substantive position about the nature of cognitively significant mental content is being adopted here. Instead, my specific contention is that there is no satisfactory construal of direct mental acquaint-

¹¹ Williamson employs the semi-technical notion *luminosity* to refine a conception of *transparency* in application to mental states. Williamson (2000, 95) defines *luminosity* as follows: A condition C is defined to be *luminous* if and only if (L) holds:

(L) For every case α , if in α C obtains, then in α one is in a position to know that C obtains.

He then advances an argument to show that no non-trivial condition is luminous. While the strategy is an interesting one, I take a different route (albeit one that is consistent with and intersects with Williamson’s strategy). Since luminosity is an explicitly *reliabilist* property, by Williamson’s own admission (Williamson 2000, 95f), it is not clear his anti-luminosity argument directly confronts any *internalist* infallibilist position that jettisons reliability constraints on rational justification/knowledge (e.g., see McGrew et al. 2007, 132ff).

ance where a mental state one is *directly* acquainted with can serve as the *infallible justifier* of a Cogito-like proposition or any other proposition about the self of significant cognitive interest.

5. Semantic infelicity

A final reply to our dismissal of infallibilism about the self calls into question our position via a revisionary maneuver.¹² This proposal tries in effect to *intrinsically* guarantee infallibility about the self, typically, by enlisting some kind of semantic-error theory of the general kind recommended by Wittgenstein (1969, 10ff). On this proposal, the putative fallibility of Cogito-like propositions is the result of some variety of semantic or conceptual infelicity. The apparent fallibility of such propositions, it is urged, is a product of linguistic misappropriation that involves the violation of conventions governing concepts about the mental. Based on the way expressions such as *I*, *think*, *doubt*, *exist*, *thinking thing* and *doubting thing* are used, the semantic error theorist reasons, it is rationally impossible to dissent from (Cogito) or (Dubito). Such dissent, it is maintained, constitutes an infringement of de facto rules concerning the proper use of language for rational subjects (where ‘rational’ itself is interpreted in some pragmatic manner). For instance, for Wittgenstein (1969) it would involve breaching implicit rules constraining the use of discrete fragments of language, i.e., making incorrect moves in what are dubbed ‘language games’. For others such as Davidson (1984; 1987) it would involve assenting to propositions (e.g., $\sim\langle I \text{ exist as a thinking thing} \rangle$) that conflict with one’s overall rational commitments,¹³ thereby generating some kind of doxastic contradiction.¹⁴ For those with Kantian inclinations it would involve endorsing proposi-

¹² This section draws on previous work of mine (see Hoffmann 2011, 245-246). The difference is that in the earlier article I consider a revisionary reply to *analytic a priori infallibilism* instead of *synthetic a priori infallibilism*.

¹³ Davidson, though, is not a proponent of semantic-error theory, nor does he profess the infallibility (as it has been defined here) of (Cogito), (Dubito), or any other proposition about the self. So it is doubly misleading to characterize him as a champion of the semantic infelicity defense of infallibilism about the self.

¹⁴ In recent work, Burge (1996) adopts a similar strategy to Davidson, one that significantly departs from Burge’s earlier defense of infallibilism about the self (1986; 1988).

tions that conflict with the postulates of Kant's transcendental deduction of the unity of apperception in consciousness, formulated in the first *Critique*.

Whatever the merits of semantic-error theory and similar outlooks, it looks to be an unpromising candidate for resurrecting infallibilism about the self. Such a viewpoint is borne out by disinterring some basic implications of the semantic infelicity proposal. Suppose it turns out the apparent fallibility of Cogito-like propositions stems from semantic impropriety of the general kind under consideration. In this instance what has been shown is that it is rationally impossible, incoherent, or inconsistent (or some combination of these things) to doubt such propositions. On the surface, though, it would not have been established that (Cogito) and (Dubito) are *logically* indefeasible in the sense at issue (see also Hoffmann 2011, 245). As I've argued elsewhere (see Hoffmann 2011, 245f), the impossibility of *doubting* a proposition is equivalent to the impossibility of a proposition's *falsity* only on the condition that logical possibility is a species of some kind of epistemic possibility, i.e., where (roughly) what is logically possible is identical to the set of propositions that is, at least in principle, knowable. Now while the logical possibility \leftrightarrow epistemic possibility thesis cannot be dismissed outright, and is typically endorsed in conjunction with some kind of semantic-error theory, this scarcely seems to threaten the anti-infallibilist argument currently on offer.

As it happens, granting the apparent fallibility of (Cogito) and (Dubito) is the product of semantic impropriety, and the concomitant logical possibility \leftrightarrow epistemic possibility thesis, in actuality involves recasting the infallibilism/fallibilism debate in a way that leaves fallibilism about propositions about the self unscathed. As I've argued elsewhere (see Hoffmann 2011, 246), this basic outlook, defensible or not, forecloses on the possibility of infallibility about the self since it involves redefining the alethic concepts of truth and falsity in explicitly anti-realist terms. The contention that (Cogito) and (Dubito) can be infallibly justified becomes essentially insensible on theories that dispute realist construals of expressions whose subject terms refer to mental phenomena, and that accordingly, contest the reality of the mental. If there is no fact of the matter about our mental states independent of the way we employ mental concepts, the infallibility/fallibility of Cogito-like propositions is a question proscribed at the outset. Davidson, for his part, explicitly acknowledges this point since, as mentioned, his account of first-person authority and privileged access does not aspire to anything in the vicinity of infallibility, indefeasibility, incorrigibility, or the like. Wittgenstein similarly

recognizes this consequence of the basic outlook under consideration. The semantic view articulated in *On Certainty*, Wittgenstein concedes, rules out the possibility of certainty as the notion is traditionally understood, i.e., from something like a language-independent standpoint.

Field, on the other hand, obliquely suggests an alternative interpretation of matters. Field (2005) has recently argued in Quinean fashion that all ‘high level’ debates about the status of purported a priori knowledge, e.g., in the areas of mathematics and logic, are best treated pragmatically – via the inferential role, illocutionary force, doxastic entitlement, or some comparable pragmatic dimension of the propositions in question. One might argue, along these general lines, that if there is no interesting question concerning the objective, theory-neutral infallibility of propositions about the self (or any other a priori proposition), and if there is a compelling pragmatic defense of infallibility about the self (as some have maintained), then infallibility about the self par excellence, the only a priori infallibility to aspire to, *is* within reach. This is correct as far as it goes but doesn’t directly confront the anti-infallibilist argument on offer. Granting the question of a priori infallibilism is a pragmatic one, and that within this framework certain kinds of a priori propositions such as those about the self are infallibly justified, only bolsters the conviction that a priori propositions about the self are not amenable to *objective* infallible justification, where *objective* is read in standard realist terms. Similar to the other semantic infelicity proposals, this one preemptively outlaws exactly the kind of infallibility about the self presently under consideration, the kind advocated by the garden variety rationalist infallibilist. Field, as it happens, is seemingly in agreement on this point (as Wittgenstein and Davidson are).

6. Concluding remarks

If I am correct, contra Descartes (1996/1641), Burge (1986; 1988), Lewis (1996) and McGrew et al. (2007), there is no proposition about the self that can be infallibly justified a priori. On the view articulated here, reason cannot furnish absolute warrant for (Cogito), (Dubito), or any other proposition about the self of significant cognitive interest. In the end, the principal defect of infallibilism about the self is found in its account of direct mental acquaintance, and in particular, in its failure to supply an account of direct mental acquaintance that enables primitive mental states to infallibly justify propositions about the self. Moreover, the class of revisionary ma-

neuers typically proposed also fail to deliver infallibilism about the self. Construing dissent from Cogito-like propositions as a form of semantic/conceptual infelicity resurrects infallibility about the self only at the unacceptable cost of redefining the terms of the infallibilist/fallibilist debate (as traditionally conceived).

The upshot of the foregoing discussion is that rationalist infallibilism and one of its two main sub-theses, synthetic a priori infallibilism, are indefensible doctrines.¹⁵ On the other hand, the arguments advanced in this paper have no direct bearing on any *non-rationalist* infallibilist doctrine, nor, as far as I can see, do they provide anything resembling a template for constructing a general argument against all forms of infallibilism. Moreover, the position defended here clearly leaves open the possibility that some species of *fallible* a priori justification may be within reach since no real reason has been given for rejecting any fallibilist model of a priori justification.

References

- BEALER, G. (2000): A Theory of the A Priori. *Pacific Philosophical Quarterly* 81, 1-30.
- BOGHOSSIAN, P. (1989): Content and Self-Knowledge. *Philosophical Topics* 17, 5-26.
- BONJOUR, L. (1996): Plantinga on Knowledge and Proper Function. In: Kvanvig, J. (ed.): *Warrant in Contemporary Epistemology: Essays in honor of Plantinga's Theory of Knowledge*. Lanham, MD: Rowman & Littlefield.
- BONJOUR, L. (2005): In Defense of the A Priori. In: Steup, M. – Sosa, E. (eds.): *Contemporary Debates in Epistemology*. Malden, Mass.: Blackwell Publishing, 98-105.
- BONJOUR, L. – SOSA, E. (2003): *Epistemic Justification*. Malden, Mass.: Blackwell Press.
- BRUECKNER, A. (1997): Externalism and Memory. *Pacific Philosophical Quarterly* 78, 1-12.
- BURGE, T. (1986): Individualism and Psychology. *Philosophical Review* 45, 3-45.
- BURGE, T. (1988): Individualism and Self-Knowledge. *Journal of Philosophy* 85, 649-663.
- BURGE, T. (1996): Our Entitlement to Self-Knowledge. *Proceedings of the Aristotelian Society* 96, 91-116.
- DAVIDSON, D. (1984): First Person Authority. *Dialectica* 38, 101-112.
- DAVIDSON, D. (1986): A Coherence Theory of Truth and Knowledge. In: Lepore, E. (ed.): *Truth and Interpretation, Perspectives on the Philosophy of Donald Davidson*. Oxford: Basil Blackwell.
- DAVIDSON, D. (1987): Knowing One's Own Mind. *Proceedings and Addresses of the American Philosophical Association* 60, No. 3, 441-458.

¹⁵ In my (2011), I also argue for the implausibility of rationalist infallibilism, although along different lines. Moreover, in that work, I argue against *both* sub-theses of rationalist infallibilism, not merely synthetic a priori infallibilism.

- DESCARTES, R. (1996/1641): *Meditations on First Philosophy*. (Trans.) J. Cottingham. Cambridge: Cambridge University Press.
- FIELD, H. (2005): Recent Debates about the A Priori. In: Gendler, T. – Hawthorne, J. (eds.): *Oxford Studies in Epistemology*. New York: Oxford University Press, 69-88.
- FUMERTON, R. (2001): Classical Foundationalism. In: Depaul, M. (ed.): *Resurrecting Old-Fashioned Foundationalism*. Lanham: Rowman and Littlefield.
- GOLDBERG, S. (1997): Self-Ascription, Self-Knowledge, and the Memory Argument. *Analysis* 57, 211-219.
- HEIL, J. (1988): Privileged Access. *Mind* 98, 238-251.
- HOFFMANN, G. (2011): Two Kinds of A Priori Infallibility. *Synthese* 181, 241-253.
- HOFFMANN, G. (2012): Infallible A Priori Self-Justifying Propositions. *Croatian Journal of Philosophy* 12, No. 1, 55-68.
- LEWIS, D. (1996): Elusive Knowledge. *Australasian Journal of Philosophy* 74, No. 4, 549-567.
- LUDLOW, P. (1999): First Person Authority and Memory. In: De Caro, M. (ed.): *Interpretations and Causes: New Perspectives on Donald Davidson's Philosophy*. Dordrecht: Kluwer Academic Publishers.
- MACDONALD, C. (1998): Externalism and Authoritative Self-Knowledge. In: Wright, C. – Smith, B. – MacDonald, C. (eds.): *Knowing Our Own Minds*. Oxford: Oxford University Press.
- MCGREW, T. (2003): A Defense of Classical Foundationalism. In: Pojman, L. (ed.): *The Theory of Knowledge*. 3rd Edition. Belmont: Wadsworth.
- MCGREW, T. – MCGREW, L. (2007): *Internalism and Epistemology: The Architecture of Reason*. New York: Routledge.
- PRICE, H. H. (1953): *Thinking and Experience*. Cambridge, Mass.: Harvard University Press.
- TURRI, J. (2011): Contingent A Priori Knowledge. *Philosophy and Phenomenological Research* 83, No. 2, 327-344.
- WILLIAMSON, T. (2000): *Knowledge and its Limits*. New York: Clarendon Press.
- WITTGENSTEIN, L. (1969): *On Certainty*. Anscombe, E. – Von Wright, G. H. (eds.). Oxford: Basil Blackwell.