

Chiu Yui Plato Tse*

Transcendental Idealism and the Self-Knowledge Premise

<https://doi.org/10.1515/jtph-2019-0014>

Published online March 6, 2020

Abstract: The relation between transcendental idealism and philosophical naturalism awaits more careful determination, i. e. whether the issue of their compatibility hinges on their ontological view on the relation between physical and mental phenomena (i. e. whether it is supervenience or emergence) or on their epistemological view on our access to mental content. The aim of this paper is to identify a tension between transcendental idealism and philosophical naturalism, which lies not in their ontological view on the nature of substances, but in their epistemological view on the relation between self-awareness and the first-personal access to mental content. I will first trace the (mis)understanding of transcendental idealism as Berkeleyan idealism to a misinterpretation of the self-knowledge premise in transcendental arguments. I will argue that transcendental idealism is not so much concerned with grounding reality of the external world as with establishing the agential nature of the first-personal perspective of experience, and it has an important implication on the meaning and function of self-awareness in transcendental idealism.

Keywords: transcendental idealism, self-knowledge, Fichte, externalism

In much of contemporary Anglo-American discussions, “idealism” has been an undesirable label reserved mostly for criticism; seldom would an academic philosopher willingly adopt the designated view (whatever it is) and call herself an idealist. Idealism is taken to imply the ontological view that reality is mind-dependent, that the world is somehow constituted by our ideas. The reason for the prevalent denial of, and reluctance toward, idealism is probably that it does not go well with the current dominant philosophical ideology of naturalism, the view that science tells us what exists and what kind of methodology is admissible. The alleged idealist thesis that what exists is mind-dependent is then obviously not what science tells us. Given the credibility of natural sciences, a

*Corresponding author: Chiu Yui Plato Tse, Dr. phil., Department of Philosophy, National Chengchi University, No. 64, Sec. 2, Zhinan Road, Wenshan District, 116 Taipei City, Taipei, 11605 Taiwan, E-mail: platotse@nccu.edu.tw

nonconformity to naturalism yields for all kinds of idealism a negative impression, including for transcendental idealism. This has contributed to the emergence of the common attitude that transcendental arguments better fare without transcendental idealism. Herein a particular understanding of the doctrine of transcendental idealism (and of Kant) is assumed, but it should not be forgotten that transcendental idealism is in fact a historical generic term covering various attempts and approaches other than Kant's (e. g. the German idealists, the Neo-Kantians, and arguably also phenomenology). The relation between transcendental idealism and philosophical naturalism awaits more careful determination, i. e. whether the issue of their compatibility hinges on their ontological view on the relation between physical and mental phenomena (i. e. whether it is supervenience or emergence) or on their epistemological view on our access to mental content. The aim of this paper is to identify a tension between transcendental idealism and philosophical naturalism, which lies not in their ontological view on the nature of substances, but in their epistemological view on the relation between self-awareness and the first-personal access to mental content. I will first trace the (mis)understanding of transcendental idealism as Berkeleyan idealism to a misinterpretation of the self-knowledge premise in transcendental arguments. I will argue that transcendental idealism is not so much concerned with grounding reality of the external world as with establishing the agential nature of the first-personal perspective of experience, and it has an important implication on the meaning and function of self-awareness in transcendental idealism.

1 A Fork in the Road from Self-Knowledge

Contemporary discussions of transcendental arguments usually give them the logical form of modus ponens: $p, p \rightarrow q/q$.¹ While this is certainly an oversimplification of the complexity and diversity of transcendental arguments and might even arguably generate the impression that the strength of transcendental arguments derives from the logical validity of modus ponens, it nevertheless brings forth one of the central features of transcendental arguments – that q is a *necessary* (and nontrivial) condition for p or, equivalently, p is a sufficient condition of q . The conditional, if true, warrants that the fact that p obtains implies the fact that q obtains as well. Concerning the nature of the premises and conclusion of transcendental arguments, there are other additional, non-logical features. As a dialectical device to defuse justificatory

¹ Similar characterizations can be found in Cassam (1987), Walker (1978), Harrison (1989).

scepticism,² transcendental arguments engage with particular sceptical positions concerning our particular beliefs, e. g. belief in external world, belief in other minds, belief in free will etc. Since transcendental arguments seek to defend for these beliefs, the belief content disputed by the sceptics shall figure in the conclusion of the arguments, instantiating q . For this device to start off, there has to be some relevant, primitive fact on which both the protagonist and antagonist can agree. Such is the nature of the factual premise that instantiates p . The second premise states a necessary condition for the possibility of the first factual premise, and this can be expressed as a conditional statement, which instantiates $p \rightarrow q$ in the schema. Using the conditional statement, the plausibility of the first unproblematic premise can then be transferred to an otherwise problematic statement, and the transcendental arguer achieves his goal through a justified inference.

In this justificatory procedure, the first premise cannot be any statement taken from common sense, for then there would be no need to argue for the belief being disputed by the sceptic, for example the belief in the existence of the external world (which is certainly part of our common sense), and the procedure would defeat the purpose of justification. Only a particular class of statements can serve the purpose, their content needs to be immediately accessible and stands in a grounding relation to something else. Epistemically, first-personal thoughts that ascribe experiential content to oneself are allegedly immune to errors through misidentification, and thus they establish a certain point of departure. Whereas the categorical premise owes its apriority to its first-personal possessive character, the conditional premise is knowable *a priori* too, “because it is, either obviously or tacitly, analytic, or conceptual, or criterial.”³ Leveraging these characteristics of the premises, transcendental arguments are taken to be making a transition from the Subjective (instantiated by a psychological fact) to the Objective (instantiated by a non-psychological fact); and if successful, transcendental arguments would then allow us to make an inference from self-knowledge to object-knowledge. Much of the following will be a closer explication of the nature of the self-knowledge premise.

Formulated in a way general enough for multifarious instantiation, the self-knowledge premise states that “We have comprehensible experience.”⁴ Two initial observations can be made at this point. Firstly, the kind of experience reported in the statement needs to be one that involves objectivity (i. e. the

² Justificatory scepticism, according to Robert Stern, is the form of scepticism that “denies that we have ‘justified opinions’ or rationally held beliefs.” See Stern (2000, p. 15).

³ Bell (2003, p. 191).

⁴ This is also Stern’s formulation in Stern (2000).

Kantian sense of experience), and the Lockean sense of experience as mere sensations would not qualify as a suitable candidate. Secondly, the reason that this premise constitutes an indubitable point of departure for transcendental argument is that it is predicated on the special epistemic access granted by the first-personal perspective. To be sure, whether there are mind-independent physical entities our experience is about is open to doubt, but the fact that, in such experience, we *judge* that there are is not open to doubt (more on this in section 3). Observably, in the contemporary discussions regarding the self-knowledge premise in transcendental argument two general approaches can be identified: either it is dislodged from the doctrine of transcendental idealism and serves only to introduce conceptual conditions that the world contingently satisfies, or it is rendered a thesis of Berkeleyan idealism, which sets constraint on how the world *must* be.

The self-knowledge premise seems *prima facie* independent of transcendental idealism, since it is a descriptive report and can in principle go with any explanatory account. The separation of idealism from transcendental arguments was first initiated by P. F. Strawson. In *The Bounds of Sense*, a reconstructive interpretation of Kant's *Critique of Pure Reason*, Strawson conceived the doctrines of transcendental idealism as "the chief obstacles to a sympathetic understanding of the *Critique*".⁵ Strawson took transcendental idealism to be the position that "the whole world of Nature is merely appearance",⁶ that mind is not only a receiving and ordering apparatus, but it also produces Nature out of the unknowable reality of things as they are in themselves. The principle of significance, which sets up the boundary of intelligibility for us, requires that empirical content be given under the conditions of our forms of sensibility – space and time. It imposes the epistemic restriction that we cannot have any intelligible cognitions of things as they are in themselves. In face of this constraint, the problem with transcendental idealism is that it involves an incoherence by traversing the bounds of intelligibility, because, in trying to limit the scope of intelligible cognition to the spatial-temporal realm, it makes a claim about the *noumenal* (and that is, non-spatiotemporal) self that it really does appear temporally.⁷ However, if this claim describes an atemporal fact, then it is unintelligible.⁸

Strawson's work stimulated not only re-readings of Kant and responses to Strawson's own position, but it also initiated a more liberal use of the terminology

⁵ Strawson (2007, p. 22).

⁶ Strawson (2007), p. 21.

⁷ Strawson (2007), p. 39.

⁸ The whole talk of affection of thing-in-themselves, which are not in space and time, on our senses would be unintelligible too.

of “transcendental arguments” in analytic philosophy, such that various forms of transcendental arguments are attributed to prominent contemporary philosophers who are not Kantian in a narrow sense (e. g. Wittgenstein, Davidson, Putnam, Searle etc.). Under the influence of Strawson, transcendental arguments have often been construed as anti-sceptical arguments purported to prove the existence of objects or at least to demonstrate the necessity of believing them to exist: they are reactionary to external-world scepticism. Cast in this light, it is commonly assumed that transcendental arguments need not involve transcendental idealism; and when it is involved, it is deemed as a dialectical (and, to many, undesirable) means to cancel the divide between the Subjective and the Objective realm.⁹

These reactions, having in view mainly Kant’s determination of transcendental idealism contained in *Transcendental Dialectics* particularly,¹⁰ tend to understand transcendental idealism as akin to Berkeley’s idealism or phenomenalism. This is problematic on at least three grounds. Firstly, transcendental idealism is not solely Kant’s invention, there is a line of post-Kantian development that substantively modifies the doctrine, hence there is no justifiable reason to confine ourselves to Kant’s particular formulation of it. Secondly, such understanding of transcendental idealism fails to link with the categorical self-knowledge premise. While most construals of transcendental arguments admit the appeal to first-personal testimony of experience as part of the premises, none had considered transcendental idealism as an account that exposes the possibility conditions for *a priori* self-cognition.¹¹ Thirdly, in the reluctant attitude towards transcendental idealism, it is likely that a confusion of two theses is involved, and they need to be separated in our analysis:

- A1. Experiential phenomena as appearances are nothing more than subjective states (I will refer to this as the *reduction thesis*); and
- A2. The cognitive capacities of rational agents constitute the appearance of experiential phenomena *as* experiential phenomena. (I will refer to this as the *constitution thesis*.)

A1 and A2 are not equivalent and they involve different commitments. Commitment to the reduction thesis has a *deflationary* impact insofar as it claims for the mind-dependence of all possible objects of experience by invalidating their

⁹ Stern (2000, pp. 49–58).

¹⁰ Kant (1998, A490–491/B518–519).

¹¹ In fact, the early development of post-Kantian philosophy, especially Reinhold’s *Elementarphilosophie* and Fichte’s *Wissenschaftslehre*, which arose in response to the skepticisms of Schulze and Maimon, produced versions of transcendental idealism that sought to supplement the hidden premises for Kant’s critical philosophy.

independent mode of existence. Commitment to the constitution thesis is not deflationary, for ascription of a constitutive role to the structure of rational agency does not necessarily deflate the mode of existence of possible objects of experience. The reduction thesis is an ontological thesis, and the idealism based on it is ontological idealism – a contrary to philosophical naturalism. The constitution thesis is an epistemological thesis, and the idealism based on it is justificatory idealism, the view that the conditions of experiential objects being cognized as such are to be derived from rational agency – also a contrary to philosophical naturalism, but for quite different reasons. If transcendental idealism is taken to be the view that the world of appearances is constituted by our subjective states, i. e. perceptions, concepts, and ideas, then the reduction thesis and the constitution thesis are likely to be confounded. Without this distinction, the opponent might refute transcendental idealism simply by refuting the reduction thesis; but with this distinction, it should be obvious that it is wrong to think that refutation of the reduction thesis implies refutation of the constitution thesis or of transcendental idealism. This confusion is exemplified by Ross Harrison’s article “Transcendental Arguments and Idealism”,¹² where he argues that transcendental arguments need not be committed to Kant’s transcendental idealism.

Harrison reconstructs transcendental arguments as having the argument form of *modus ponens* as well. Idealistic transcendental arguments attempt to infer the actuality of a disputed claim from its necessity. If causal order of the natural world is necessary, then it is actual; if existence of the external world is necessary, then sceptical doubt of it is untenable. This line of reasoning is fairly straightforward and can formally be thus represented: $\Box p/p$. To complete the standard components of transcendental argument as suggested earlier, a factual premise and a conditional premise need to be added and the idealistic argument form, according to Harrison, would be:

1. $\Box j$
2. $\Box(j \rightarrow p)$
3. $\Box p$ (from 1, 2)
4. p (from 3)

The harmless assumption j states that there is judgement. Conjoined with a necessity operator, Harrison interprets $\Box j$ to be the assumption of idealism, which says that what there is is necessarily judged – *esse est concipi*. On this construal, transcendental idealism and Berkeleyan idealism are identical. However, this interpretation of $\Box j$ is problematic, for it confuses two reading of $\Box j$:

¹² Harrison (1982).

- B1. Necessarily, there is judgement. (I will refer to this as the *self-knowledge premise*.)
- B2. What there is is necessarily judged. (I will refer to this as the *Berkeleyan premise*.)

$\Box j$ should be read as B1 and not as B2, this confusion misled Harrison to the conviction that transcendental arguments should do away with transcendental idealism. But the self-knowledge premise and the Berkeleyan premise are not equivalent. Although both are modal statements, the Berkeleyan premise that what there is must be conceived is a *de re* modal claim, whereas the necessity in the self-knowledge premise is *de dicto*. The former makes a non-trivial, universal claim about the entire admissible ontological realm, and it should now be obvious why it is problematic for transcendental arguments. Recall our earlier determination of the nature of transcendental arguments – they are purported to establish anti-sceptical conclusions. For that purpose, the dialectic of transcendental arguments must respect two constraints: 1. it needs to start with a premise that even sceptics have no qualms about, so it must not assume what has to be proven; 2. neither should it concede too much to scepticism, otherwise it would not be establishing anti-sceptical conclusions. Should $\Box j$ be read as B2, in other words, should transcendental idealism be committed to the Berkeleyan premise, then indeed it would be rightful to exclude the premise from transcendental arguments, as it does not abide by the constraints. The resulting problem would not only be that, as Robert Stern took it,¹³ the transcendental idealist joined the sceptics in denouncing the objectivity of the external world and the Berkeleyan premise defeated the purpose of transcendental arguments, but it also begs the question against the sceptic: how does one know that things are as the Berkeleyan premise says they are? If the Berkeleyan premise does not seem to be of much use in resisting scepticism or is itself subject to sceptical doubt, then the immediate consequence to be drawn, given that we have just sorted out two nonidentical readings of the premise of transcendental idealism, is that we should not accept the B2 reading of transcendental idealism.

What about B1? The self-knowledge premise appears to be an existential claim about an entity-type called “judgement”, but it needs not and should not be taken in such literal manner. There is no such *thing* as judgement except as a result of the act of judging. A unique relation holds between the act of judging and the fact that there is judgement. Epistemically, for a subject *S* engaged in the act of judging that *P*, “I judge that *P*” is accessible *a priori*. Alethically, for a subject *S* who judges on the truth of the statement “I judge that *P*”, the very act of judging makes true that

¹³ Stern (2000, p. 49).

statement. Under the first reading of \Box_j as a self-knowledge premise, the necessity attached to the self-ascriptive judgement points to the special epistemic-alethic status grounded in the first-personal character of experience. How exactly is this grounding relation to be understood? This question will be tackled in the next two sections. For now, I believe the above analysis has sufficiently shown that the problems that beset the Berkeleyan premise do not amount to a decisive refutation of transcendental idealism, but should rather be taken as guidance on how (not) to understand the thesis of transcendental idealism. Instead of a reductive and deflationary account of reality as that of Berkeley's, it offers the necessary conditions of the possibility of self-knowledge, and as such its role would be better appreciated if we place it within self-directed instead of world-directed transcendental arguments.¹⁴

2 Non-Introspective Self-Consciousness in Transcendental Idealism: The Case of Fichte

The result of the last section suggests that there is another strand of transcendental idealism that holds the constitution thesis (and not the reduction thesis) without arguing from the Berkeleyan premise. In this section, I would like to suggest further that this other strand of transcendental idealism has another purpose to serve other than refuting external world scepticism, though it also delivers an indirect argument in response to it. This alternative conception of transcendental idealism, like the strand critically exposed under post-Strawsonian light, has its historical roots as well. Unless one only wants to stipulate the definition of "transcendental idealism" to refer to the doctrines contained only in the sections of "Transcendental Deduction" (for which there are significant differences in the two editions of *Critique of Pure Reason*) and "The Refutation of Idealism" (which was added only in the second edition) in Kant's First Critique, there is no principled reason to look away from modified versions of transcendental idealism that proclaimed to improve on Kant's account. While there are numerous attempts that seek to reconstruct and assess Kant's transcendental idealism under an analytic conception of transcendental arguments, similar treatment of post-Kantian idealism is scant. Paul Franks has called for our attention to the historical debates about Kant's transcendental method in the late eighteenth-century, which interestingly anticipated the contemporary debates in the twentieth-century concerning the success of transcendental arguments in

¹⁴ I am following Cassam's (2003) use of the term "self-directed transcendental arguments".

fending off scepticism.¹⁵ Franks' treatment of Gottlob Ernst Schulze's sceptical criticism of Carl Leonhard Reinhold ingeniously connects a historical controversy in the 1790s with Stroud's sceptical challenge, and thereby brings out the relevance and originality of Fichte's strategy in response to the so-called reality and uniqueness problem raised by Schulze.¹⁶ Critical revisions and transformations of Kant's transcendental idealism began shortly after the publication of the First Critique, so the talk of an alternative conception established in the last section does not merely point to a theoretical possibility, but at the same time also a historical actuality that contemporary discussions of transcendental argument tend to ignore. (Having that said, the following is not a recapitulation of Franks' account.)

The alternative strand of transcendental idealism, which can presumably be traced to Reinhold's proposal of *Elementarphilosophie* and reaches a more mature shape in Fichte's *Wissenschaftslehre*, does not take issue with external world scepticism directly; instead, it grapples with a more eminent issue, which could well be called internal capacity scepticism. Since the task of Fichte's *Wissenschaftslehre* is to show that experience is fundamentally constituted by the self-determinative nature of reason, which is easily a target of scepticism, his idealism argues from the self-knowledge premise to the efficacy of reason via a transcendental exposition of the structure of self-consciousness. Here, an important difference between self-knowledge and self-consciousness should be noted. Whereas the former is an empirical report from an individual first-personal perspective, the latter points to an abstract structure that underlies all empirical first-personal thoughts and conditions object consciousness. Similarly, there is an important difference between the individual I and the absolute I. The former is a person, whose bodily states and mental episodes are the contingent objects of her introspective self-cognition; the latter does not refer to any person or spatiotemporal thinking substance, it designates the self-referential, autonomous, and perspectival nature of reason that cannot be abstracted away without forsaking intelligibility. Arguing from the former to the latter, transcendental idealism could ground the special epistemic status of first-personal thoughts in the universal structure of self-consciousness. One could, I suggest, understand Reinhold's manoeuvre of

¹⁵ Franks (2003).

¹⁶ I believe Franks is being over-cautious in stressing the need to "avoid an anachronistic imposition of contemporary assumptions" (2003, 120) in examining the species of transcendental arguments proposed by the post-Kantians. This hermeneutical caution is harmless and, in usual circumstances, well justified. But if anachronism is a genuine issue, then it must be tackled when one examines Kant's transcendental arguments. Yet, no one seems to hold such attitude regarding Kant, so it is uncalled for to raise this issue in examining the post-Kantians' transcendental arguments.

regressing from the fact of experience to the structure of the faculty of representation as serving the same purpose; and one of the reasons that he fails is that the empirical and the transcendental level of justification was not adequately distinguished.

Reinhold's philosophy of the elements (*Elementarphilosophie*) is an attempt to systematize Kant's transcendental philosophy by providing for it a fundamental principle of the ultimate foundation. He identifies representation (*Vorstellung*) as the common denominator underlying intuition, concept, and ideas and he sees the faculty of representation as the single most important faculty of mind. His principle of consciousness (*Satz des Bewusstseins*) states that: "In consciousness representation is distinguished by the subject from the subject and object, and is related to both". Formulating the fundamental principle in this way, his epistemology proposes a triangulation of the representing subject, the represented object and the representation *per se* and provides a shorter path to transcendental idealism, one that diverts from Kant's version and incurs its own sceptical criticisms.

In 1792, Schulze's sceptical review of Reinhold's theory of the human faculty of cognition and Kant's critical philosophy appeared anonymously with the title "*Aenesidmus Or Concerning the Foundations of The Philosophy of The Elements Issued by Prof. Reinhold in Jena Together with a Defence of Scepticism Against the Pretensions of The Critique of Reason*". In *Aenesidemus*, roughly speaking, two major charges against (Kant's as well as Reinhold's) transcendental idealism were mounted and they can be formulated in the following way: C1. Transcendental idealism, in its inference by the principle of sufficient reason, makes an illegitimate move from a proposition about thought (viz. "representations") to a proposition about being (viz. "things-in-themselves");¹⁷ C2. What transcendental idealism takes to be the necessary condition for the possibility of experience might turn out to be a contingent limitation of human culture and hence not a genuinely necessary condition of possibility.¹⁸ Since intuition for Kant and Reinhold can only be sensible, there can be no pure intuition of the mind or the faculty of representation as the determining ground for synthetic judgements *a priori* or representations. Given this epistemic constraint, the first charge raises the question of what entitles one to claim for the mind-in-itself the status of an actual cause, and that is, its reality. The second charge calls the *a priori* status of the principle of consciousness into question. Although the principle is taken to describe a fundamental structure of representation, the reflection involved is based on empirical self-observation and must therefore be limited to empirical representations. Since in this light the proposition of consciousness expresses only an empirical fact, the sceptic can

¹⁷ Schulze (1985/1792), pp. 108–09, 114–15).

¹⁸ Schulze (1985/1792), pp. 117–118.

easily deny its necessity by allowing experiences that contradict it. This would mean that the principle of consciousness, which is supposed to give us the necessary conditions for the fact of experience, is not unique – clearly an undesirable consequence for Reinhold. These charges tie well into the naturalist view that the reality of the mind is dubitable (because mental causation is dubitable) and the supposed epistemic specialness of thought content is only read into the mind by us. Fichte’s enthusiasm for uncovering the foundation and establishing the systematicity of transcendental idealism was motivated by the need to meet the reality and uniqueness problem,¹⁹ and his emphasis on the first-person point of view and the progressive development of the necessary acts that constitute objectivity should be understood within this context.

Fichte’s emphasis on the first-personal point of view in his transcendental reflection is part of his antidote to the reality problem and the uniqueness problem. In his “Review of Aenesidemus” (1794), Schulze’s criticism of Reinhold brought Fichte, who for a short time was a proponent of Reinhold’s project, to concede that the principle of consciousness is based on an empirical self-observation and abstraction. To avoid the problem of recurring to empirical fact in the transcendental claim, Fichte turns to the original act (*Tathandlung*) in contrast to fact (*Tatsache*),²⁰ a transcendental feature that characterizes the I without enmeshing any dogmatic conception of facticity. The sense of self-consciousness as the original act impedes us from taking it as introspection, and points to a non-introspective notion of self-consciousness. But what function does this non-introspective notion of self-consciousness play in Fichte’s idealism?

The introduction of “*Tathandlung*” in transcendental argument serves two purposes. Firstly, it forestalls the alternative of explaining conscious experience as the effect of the affection caused by things-in-themselves.²¹ Secondly, it postulates the spontaneous nature and freedom of the intellect as a necessary condition of conscious experience, a condition which physicalist denies.²² In section 1 of *Foundation of the Entire Wissenschaftslehre* (1794/95), Fichte establishes the absolute, and that means ineliminable and unconditional, status of the I by showing the absolute necessity of the self-positing act as the primary unity-conferring

19 For historical precision, we should note that Fichte also takes himself as answering an analogous form of scepticism raised by Salomon Maimon. Maimon’s attack on the applicability of the *a priori* concepts of the understanding on the *a posteriori* given content from sensibility challenges the legitimacy of schematism in a fashion similar to semantic externalism’s challenge to individualism.

20 Fichte (1964–2012: I, 2, p. 46).

21 It is in fact conceptually questionable whether an argument that has thing-in-themselves do the most crucial work can ever qualify as a transcendental argument.

22 For a recent reconstruction of Fichte’s critique of physicalism, see Tse (2018).

activity, the vehicle of all representations, without which no content of cognition can be intelligible. To demonstrate this, he reflects on the uncontroversial logical law of identity through the proposition $A = A$ and argues that the law of identity is true not in virtue of the existence or property of objects but in virtue of the I that posits itself. The law of identity says that, in a more precise manner of predicate logic, $\forall x (x = x)$. What makes it a law of *thought*? From the standpoint of ordinary consciousness, it seems more like a law of object, according to which whatever that exists must have the property of being identical to itself. But, from the standpoint of philosophical reflection, for these strings of symbols to play any cognitive role, they must *mean* or be *interpreted* in a certain way. If 'tree' is taken to refer to trees on one occasion, to apples on another, and to other random things on yet other occasions, then we do not have a proper interpretation and hence no truth assignment. So, once 'A' is assigned a particular meaning in a model, it must be interpreted consistently; the interpretation function of a model requires that the meaning of terms be identical in that model. Under this light, the reason of Fichte's peculiar translation of the self-identity statement into a conditional form, "*if A exists, so A exists*",²³ is to dislodge an unconditional epistemic requirement from ontological commitment. Fichte points to the first unconditional epistemic requirement for intentional content by analysing the unitary function expressed by the conditional positing. The conditional-sounding connective of 'if-so' serves then to make explicit the rule that the occurrences of a representation, if it is to be a unit in our intelligible thought, must be brought together under a unified intention. Since the identity of intentionality is not grounded on the existence of objects, it can only be grounded on the nature of the subject that is capable of entertaining intentions. It is in this light, I suggest, that Fichte's claim that "X is really posited in the I" can be best understood.²⁴ This nominalized I, under which the necessary connection X must be posited, is not an ordinary, specific individual, but rather a generic structure exemplified by all finite rational agents who possess the first-personal concept of self.²⁵

As sure as the self-reverting fact/act has a function that is necessary for the identity and intelligibility of intention, the object of representation must be structurally distinct from that act. With the second axiom in section 1, Fichte

²³ Fichte (1964–2012: I, 2, p. 257).

²⁴ Fichte (1964–2012: I, 2, p. 257).

²⁵ For a concise exposition of the intricate relation between the absolute I and the ordinary I, see Zöllner (2007, pp. 253–255) Zöllner rightly observes that the self-reverting activity of the absolute I does not amount to full-blown, manifest self-consciousness, it functions rather "as the closest transcendental ancestor to this phenomenon." Very roughly speaking, my distinction between non-introspective self-consciousness and introspective self-consciousness cuts similarly, the former is a transcendental ancestor to, and a necessary condition of, the latter.

introduces the Not-I, which, in opposition to the I, is essentially not self-reverting. If the *Tathandlung* is a transcendental enabling condition that makes introspective self-knowledge possible, then the Not-I is a transcendental disabling condition that introduces a heterogeneous element in the structure of representation. For the sake of the unity of consciousness that is universally shared by finite human beings, the I and the Not-I must cohere, however mutually exclusive the first two principles might appear to be. Hence, the third axiom states that the finite I and the finite Not-I must be op-posed within the I. The analysis that explicates the transcendental activities involved leads to the bifurcation of a theoretical part and a practical part in the *Foundation of the Entire Wissenschaftslehre*, it will far exceed the scope of this article to pursue the details. Nevertheless, it is worth remarking on the anti-dogmatic and anti-naturalistic nature of the I drawing on some of the results from the relevant parts.

Fichte's account of the absolute I, which is above all self-reverting, is at the same time an account of self-consciousness that is not an instance, but rather a determining condition, of ordinary consciousness and self-cognition. The explication of the I's necessity for content identity just sketched above is clearly anti-naturalistic, as it does not invoke any naturalistic data or procedure in the explication. While Fichte's presentation of axioms (*Grundsätze*) and theorems (*Lehrsätze*) might give a rationalistic impression, it is unambiguously guided by a line of transcendental reflections. The primacy of the absolute I is not dogmatically assigned; instead, it is primary because of its epistemic immediacy and self-verifying nature. This lays down the non-natural foundation of the indispensable role of self-consciousness as a transcendental condition for the possibility of object consciousness. As Fichte further develops the inner structure of the I through a series of oppositions and gives it more substance in practical terms, it turns out that the original act consists in reason's activity of *self-determination*. On the one hand, his conception of the unity of consciousness is akin to Kant's in that the synthetic function takes priority over the analytic. For Kant, the analytical unity of apperception is only possible on the ground that the mind "can combine a manifold of given representations *in one consciousness*;"²⁶ whereas for Fichte, the unity of consciousness hinges on a fundamental synthesis of the I and the Not-I within the I.²⁷ On the other hand, Fichte recognizes the unity of consciousness as an ideal and not as a brute given fact (while the *possibility* of the unity of consciousness must be actual as a fact), because the fundamental synthesis, which is a necessary condition for the possibility of experience, is an infinite task (*Aufgabe*) of determination. Fichte's transcendental reflection of it reveals theoretical as well as

²⁶ Kant (1998, B133-B136).

²⁷ Fichte (1964–2012, I, 2, p. 275).

practical significance. In the latter aspect, practical reason takes up a much broader function than it has in moral philosophy and is also brought into the explanation of the possibility of knowledge. As a consequence of the structure of the synthetic unity of consciousness that Fichte describes, striving becomes a necessary expression of the *Tathandlung* of the absolute I in the face of a Not-I, it is a transcendental feature in all the activities that in one way or another try to subjugate a non-conceptual element under the horizon of intelligibility. The self-determining activity of reason has then a regulative function, by which the various capacities of human mind are reciprocated and organized.

Given the anti-dogmatic status of the I, it would be a misunderstanding to apply the reality problem on Fichte's original theory of the I; its self-determination and object determination is not grounded on the same sort of causal efficacy that governs the substantial exchange of properties. Although Fichte occasionally characterizes the absolute I as "the I-in-itself",²⁸ it is clear that the I is not a thing nor a substance, and its reality is coextensive with its ideality. A purposive act is real insofar as it strives for an ideal, and it is ideal insofar as it is driven to strive. The reality of the activity does not depend on whether that ideal is realized, an activity that falls short of its goal is not any less real. Fichte's transcendental idealism faces up to the uniqueness problem by grounding the objectivity of representations without Kant's epistemological dualism (namely the rigid distinctions between understanding and sensibility, appearance and things in themselves). To demonstrate that reason's free act of self-determination is a necessary condition of experience, transcendental idealism, in Fichte's own words, needs to "deduce the system of the necessary modes of operation, and with it concurrently the objective presentations created thereby, from the fundamental laws of the intellect, and so allow the whole compass of our presentations to come gradually into being before the eyes of its readers or listeners."²⁹ The "deduction" at issue here is certainly not one of logical deduction. For Fichte, the procedure of this transcendental deduction is nothing other than exposing all the necessary conditions involved once one accepts the fundamental principles of reason.³⁰ Yet, this path of transcendental idealism involves a profound transformation of Kant's theory of capacity.

Fichte's explication of the thetic act, the anti-thetic act, and the synthetic act is only the first stage that delivers the general, determinable capacity of reason. On that

²⁸ Fichte (1964–2012, I, 4, p. 190).

²⁹ Fichte (1964–2012, I, 4, p. 201; 1982, p. 25).

³⁰ See Fichte's own description of the procedure of deduction (1964–2012, I, 4, p. 201; 1982, p. 25): "It shows that what is first set up as fundamental principle and directly demonstrated in consciousness, is impossible unless something else occurs along with it, and that this something else is impossible unless a third something also takes place, and so on until the conditions of what was first exhibited are completely exhausted, and this latter is, with respect to its possibility, fully intelligible."

basis he deduces intuition, understanding, imagination, and the power of judgement as more determinate forms of capacity. Objects of possible experience are a result of the interplay between the determinate capacities of the intellect. In other words, object-consciousness arises from the unfolding of transcendental self-consciousness. Through a deduction of representation, Fichte offers a picture of how object-consciousness emerges as a product of co-determination by various determinate modes of activities. The determination of object-consciousness via the self-determination in self-consciousness indicates that Fichte's transcendental idealism is in the first instance self-directed and offers only an indirect argument for the existence of the external world.³¹ Since the focus of this paper is not so much on object-knowledge as on self-knowledge, I shall only sketch the skeletal framework of the interdetermination of capacities that makes representation of objects possible.

As primitive states determinate feelings (which are limitations of the subject) and striving are given to the I. Imagination is a striving force that combines determinate feelings and sensations into objects, but the power of imagination does not by itself provide the individuation conditions. So, the understanding takes up an auxiliary (but yet essential) role of halting the blind drive of imagination and setting down the boundaries of intuitive content. Intuition which delivers our awareness of an object is then derived from this fixation of imagination through the understanding. Based on the work of imagination and the exertion on intuition, the understanding delivers determinate objects to the power of judgement. The power of judgement, as the capacity to reflect on or abstract from determinate object, rises to the notion of *determinable* object, which lays the conceptual foundation for the categories of causality and substantiality. The entire interplay between the determinate capacities just described is conditioned by the self-determination of reason qua the purposeful capacity striving for autonomy and inner harmony. Embedded in the holistic web of inter-determinations for the sake of self-determination, the objectivity of representations turns out to be a necessary outcome of reason's development, and, for Fichte, "this is why an object must necessarily exist for us."³²

3 Self-Knowledge and Externalism

The relevance and significance of the first-person perspective for transcendental arguments does not go completely unnoticed,³³ but the bearing it has on the

³¹ This summary is based on a section called "Deduction of Representation" in Fichte's *Foundation of the Entire Wissenschaftslehre*, see Fichte (1964–2012, I, 2, pp. 369–384).

³² Fichte (1992, 234).

³³ See Franks (2003) and Bell (2003).

justificatory strategy of transcendental philosophy remains unclear, and peculiarly, nothing has been said in response to the problematic supposition that, if self-directed transcendental arguments are to reveal the structure of our cognitive faculties, the mind is transparent to itself. Having determined the sense of the constitution thesis with reference to the alternative, Fichtean account of non-introspective self-awareness, we are now in a position to make an overdue response. My aim in this section is to show that semantic externalism presents a difficulty to content transparency, and this will further guide us how (not) to understand transcendental idealism.

Content transparency is here understood to be an epistemic feature of mental content that allows a privileged access such that the owner of the content cannot be susceptible to systematic errors about the content. According to Paul Boghossian's formulation (with reference to Michael Dummett) of the thesis of epistemic transparency, if two token thoughts (belonging to a thinker) possess the same content or distinct contents, then the thinker must be able to know *a priori* that they do.³⁴ It seems to be a necessary condition for making reflective judgments in critical reasoning, and is for this reason also necessary for the assessments of a subject's rationality or psychological explanations of his behaviour. Embedded in the project of achieving reason's self-cognition, it seems that transcendental idealism is a viable doctrine only if the first-personal perspective entitles us a special, *a priori* access to our mental content. If this is true, then transcendental philosophy, as a project that aims at *a priori* self-cognition (that is, cognition of forms of cognition), is committed to the view that certain relevant aspects of our beliefs, desires, and perception must be transparent to us.

Certainly, it makes sense to say that we do not know ourselves or we do not know as much about ourselves as we might assume, especially when it comes to unconscious psychological mechanisms or pre-reflected stereotypical beliefs. Yet, there are also several occasions where self-intimacy is undisputed and *undisputable* if assessment or psychological explanation of an agent's behaviour is to be viable. When I am composing this sentence, not only must I know what my occurrent mental content is that is being conveyed, I must also be assumed by my audience (when I am uttering this sentence) to know the thought content I am currently trying to communicate if I am to be held and assessed as a rational agent who is responsible for whatever ingenious or mediocre content being presented. My epistemic access to my mental content is even though to be *a priori*: there is no need to conduct an empirical investigation (and indeed it would be strange to conceive one) to find out what I am thinking. Some notion of epistemic transparency has to be at work. But of course, I do not thereby know every fact there is of

³⁴ Boghossian (1994, p. 36).

the content about which I possess a representation. In asserting “I am thinking that there are physical entities” or “I judge, herewith, that there are physical entities”, I cannot be wrong even if there are in fact no physical entities. This class of statements makes up what Tyler Burge calls “cogito-like judgements”.³⁵ This type of judgements is epistemically special insofar as it is not based on any observation or perception and its thought is contextually self-verifying. The correctness of such judgements is not determined by verification of environmental conditions. It is immediate and is not susceptible to brute error as in cases of perceptual belief.

But this somewhat common-sensical view on epistemic transparency is called into question by philosophical naturalism. Presumably, philosophical naturalism implies mental content externalism (or even eliminativism in extreme case). Content externalism is the view that the mental contents of a subject are partly individuated by facts about the physical or social environment that are not transparent to the subject. This view involves a rejection of the thesis of epistemic transparency, and a consequence of the rejection is that empirical investigation is needed in order to determine what our thought content actually is. If nothing in the mind constitutes for it a special epistemic access to itself, then there is nothing to substantiate the transcendental idealist claim that the object of intuition conforms to the forms of thought, and there would be no ground for any non-empirical investigation of the structure of our cognitive faculties. In the entry “Externalism and Self-Knowledge” on *The Stanford Encyclopedia of Philosophy*, Ted Parent succinctly identifies the tension between externalism and self-knowledge as one between their major thesis:³⁶

EXT: Thought content is determined partly by the environment.

SK: A subject can know from the armchair what content her thoughts have.

If they are genuinely incompatible, what one should make out of this tension is up for interpretation. If one takes *a priori* self-cognition as primary, then the tension will appear as a challenge to the externalists; if one takes externalism to be the correct account of the determination of thought content, then the tension with it might call into question the possibility of *a priori* self-cognition. Perhaps this is the genuine place that transcendental idealism should turn to.

Semantic externalism is advanced by Putnam’s famous Twin Earth thought experiment, concluding with the dictum that “Meaning just ain’t in the head”.³⁷ While the thought experiment was not originally devised as an argument against self-knowledge or the idea of epistemic transparency; its influence on subsequent

³⁵ See Burge (1996).

³⁶ Parent (2017).

³⁷ Putnam (1975, p. 225). Content externalism has since then been the orthodox theory of meaning (though only a thin majority according to a poll by PhilPapers in 2009).

discussions in philosophy of mind is immense. Burge, whose anti-individualism about thoughts individuation was inspired by Putnam's externalism, devises a slow switching scenario which would easily lead one to conclude that EXT and SK are incompatible.³⁸ It is based on a slight modification of Putnam's original version, but considers the mental goings-on of just one subject, *S*, who is transposed between twin earths (W_1 , W_2) without knowing that such switches occur. Whereas the substance known as "water" on Earth is H_2O , the substance known as "water" on twin Earth is XYZ; otherwise the twin earths are in every way qualitatively the same such that *S* does not sense any difference in the way things feel. After a series of slow switches, causal relations of reference will have been established between *S*'s concepts of "water" and H_2O in W_1 and XYZ in W_2 respectively. Since now there are two non-overlapping causal links, with each connecting "water" to two different substances, there are two sets of individuation conditions that set the content of "water" at W_1 and the content of "water" at W_2 apart. To be sure, at each moment when *S* entertains the thought that "I judge that there is water", he knows what he is thinking about. However, if *S* is told of the switches and is asked to differentiate the concept of water that has H_2O as its content and the concept of water that has XYZ as its content, he will not be able to tell the difference.

While the externalist incompatibilist (about the joint truth of EXT and SK) might conclude from this thought experiment that "a person could not know what thoughts he had unless he undertook an empirical investigation of the environment which would bring out the environmental differences",³⁹ Burge insists nevertheless that this is absurd and stresses on a fundamental disanalogy between perceptual knowledge and the basic self-knowledge expressed by cogito-like judgements.⁴⁰ While perceptual knowledge is objective in two different senses, namely in being susceptible to brute errors and in being impersonal, self-knowledge differs from perceptual knowledge in both these aspects. Self-knowledge is not susceptible to brute errors in the way perceptual knowledge is, because there is no gap between one's thoughts and the subject matter. In making self-ascriptive judgement "I think that *p*", one does not take one's thought that *p* as a mere object, but more as a just executed task, which bears a non-contingent relation with the reflexive judgement itself. Nothing about any particular person's making the judgement is relevant to the justification of any empirical belief about the physical

38 See Burge (1988). Burge's neo-rationalist account of self-knowledge is representative of the kind of semantic externalism that preserves the special epistemological (which he also describes as "environmentally neutral") status of self-ascriptive judgements.

39 Burge (1988, p. 653).

40 For a more comprehensive survey of not just Burge's but also Davidson's approach to the post-Putnamian puzzle of the reconcilability of externalism and self-knowledge, see Bilgrami (1992).

world, but the first-person point of view matters, figures even as a necessary condition, in self-knowledge.

Burge's treatment of the slow switch thought experiment points to the epistemic specialness of self-knowledge and its simultaneous consistency with first-order beliefs that enter into the scope of self-ascribed propositional attitudes. The moral is that, despite the external determination of the content of our first-order perceptual beliefs by the environment, the determination of the content of our second-order self-reflective judgement remains individualistic. Or in Burge's own terms, the epistemic entitlement to one's own mental kinds "survives" environmental switches. In explaining this epistemic entitlement, Burge points to its role in critical reasoning. Should our thoughts and beliefs be operable in our critical reasoning and be evaluable according to rational norms, then at least some of our thoughts have to be *knowledgeably reviewable*.⁴¹

However, it is not quite enough to settle the issue. Boghossian sees likewise the tension between content externalism and self-knowledge as a genuine puzzle, but he thinks that Burge's appeal to the normative role of self-knowledge cannot safeguard its epistemic specialness, and concludes that externalism about thought content insolvably violates epistemic transparency. To show this, Boghossian also invokes the slow switching thought experiment and generates scenarios in which the subject caught in switching is making logical mistakes without being able to detect them. One of the scenarios runs like thus:⁴² Given the same switch setting, suppose Peter had met the Italian tenor Luciano Pavarotti at Lake Taupo once in his past experience on Earth, now reading news about the Twin Pavarotti on Twin Earth reminisces about the past encounter, and he takes his memory representations of the Earth Pavarotti to be relating to the Twin Pavarotti. He might be tempted to think that the following is an instance of *modus ponens*: Whoever floats on water, gets wet. Pavarotti once floated on water. Pavarotti once got wet. With this line of reasoning, Peter is inferring a conclusion about the Twin Pavarotti (that he once got wet) based on a general rule and a memory belief about the Earth Pavarotti. What seems to Peter to be a valid inference is in fact not.⁴³ The most important and relevant result is that, while there is still a sense of immediate, reflexive access to one's own occurrent beliefs preserving the immunity of errors through misidentification,⁴⁴ this access does not warrant the epistemic capacity to introspectively discriminate between

⁴¹ Burge (1996, p. 98).

⁴² Boghossian (1994, p. 44–45).

⁴³ In this case, there are two token expressions of the same syntactic type (<Pavarotti>₁ for Earth Pavarotti and <Pavarotti>₂ for Twin Pavarotti) in Peter's language of thought, but they possess different semantic values and this fact is not *a priori* accessible to Peter.

⁴⁴ It means that if Peter possesses a token belief "p", he must also possess the token belief "I believe that p".

mental contents. This inability to discriminate between mental contents makes the subject susceptible to logical errors which she cannot detect, and this insensitivity to logical errors tells against the normative function of critical reasoning and self-evaluation that Burge alleges for our reflexive, self-referential thoughts. Hence, if even parts of the mental content of self-ascriptive judgements is determined non-individualistically, then the certainty and immediacy involved in basic self-knowledge cannot warrant infallibility.

Boghossian's line of argument targets a conception of self-knowledge that incorporates epistemic transparency of the identity as well as difference of content, which apparently exceeds Burge's notion of basic self-knowledge. So, to what extent Boghossian's criticism really affects Burge's account remains an open question that cannot be tackled in this article, and I shall not delve further into the intricate implications regarding the logical properties of thought contents that variations of slow switching arguments seek to bring out. To return to the task stated at the beginning of this section, the following two questions need to be answered.

Firstly, how does Fichte's transcendental idealism stand to this issue in contemporary philosophy of mind? Pro Burge, Fichte's conception of transcendental idealism, which is to be conceived in a constitutive instead of a reductive sense, rests on a notion of self-consciousness that offers the transcendental ground of basic self-knowledge, which one gains, according to Burge, "simply by thinking it while exercising second-order, self-ascriptive powers".⁴⁵ Since basic self-knowledge claims are not logical truths, nor are they contingently true, any explanation of the relation between one's self-ascriptive powers and one's thoughts that appeals to a purely logical or empirical (or psychological) account would not be adequate. Fichte's transcendental idealism provides a theory of the transcendental activities of reason, which establishes a non-empirical and yet also non-logical relation between the self-ascriptive powers and the presentation of content.

Secondly, does transcendental idealism require transparency of the mind in Dummett's sense? If a moral for transcendental idealism has to be drawn from Boghossian's observations of the slow-switching thought experiment, it should be that the transcendental idealist, in invoking a non-empirical sense of self-consciousness as an *a priori* condition of possibility, needs to be wary of an implicit adoption or implication of the Dummettian sense of epistemic transparency, which sets too high a standard for empirical self-cognition. In other words, self-verification, self-determination, and self-illumination are separate, non-coextensive properties. This diagnosis should help us situate (and not prematurely repudiate) transcendental idealism in the face of the challenge of naturalism, it shows that a transcendental/empirical distinction should be made between what has often

⁴⁵ Burge (1988, p. 656).

been referred to as the “immunity to errors through misidentification” in self-knowledge and the infallibility of its content. The immunity to errors through misidentification pertains to the epistemic subject universally due to its self-ascriptive power, but the infallibility of content pertains only ideally to the epistemically unlimited subject. For a finite rational being, the former obtains but not the latter. While Peter may be wrong in thinking that the Pavarotti he was thinking of in W_2 is the same Pavarotti he had thought of in W_1 , it remains true that Peter bears a special relation to his thought content, such that whereas “Pavarotti once got wet” is fallible, “I judge that Pavarotti once got wet” is made true by the judging itself. Therefore, *a priori* self-cognition with fallible content is a possible species, one that has received almost no attention in the classical literature of transcendental idealism.

While there may well be passages in Fichte where his notion of apriority seems to imply indubitability, it does not necessarily indicate that Fichte subscribes to the problematic notion of infallibility, under which the finite intellect would be inherently sensitive to extra-mental individuation conditions of content. For instance, in his “Review of Aenesidemus” Fichte objects the treatment of thoughts as things-in-themselves and assumes that it does not make much sense to speak of thoughts as if they were independent of our thinking and yet somehow determined our judgements.⁴⁶ While I do not want to say that Fichte has thoroughly and unambiguously distinguished self-verification, self-determination, and self-illumination in all his claims about the apriority of self-consciousness, it is important to note that his objection was primarily a reply to Schulze’s criticism concerning the status of the philosophical claims in Kant’s and Reinhold’s theories of mind and cognition. Those claims of transcendental idealism are in the first instance about our capacities of cognizing objects, and not directly about objects as such; they are therefore self-ascriptive. In this light, I suggest, Fichte’s objection preserves the gnoseological intimacy (or the assumption of the unintelligibility of alienated content) for the transcendental, second-order enterprise. But this should not be taken to mean that first-order empirical cognitions that take objects as part of their content would become infallible simply in virtue of carrying such cognitions over to a higher-order reflection.

4 Conclusions

The conclusions drawn from the previous sections can be summarized as the following:

⁴⁶ Fichte (1964–2012, I, 2, p. 53).

1. There are unreflected reasons underlying the common reluctant attitude to transcendental idealism in contemporary discussions of transcendental arguments.
 - a. There is a confusion of the reduction thesis and the constitution thesis.
 - b. Transcendental idealism was narrowly defined on Kant's view only.
2. A misreading of the categorical premise of transcendental arguments can be identified in Harrison's attempt to dispense with transcendental idealism. The correct reading should render the categorical premise a self-knowledge premise, not a Berkeleyan one. Transcendental idealism offers the ground for the former, not for the latter.
3. An alternative transcendental idealism can be found in Fichte. The account was motivated in part by the need to avoid the sceptical attack that fell on Reinhold's approach, which delivers the conditions for the possibility of experience from an introspection on the fact of consciousness. The significance of the first-personal point of view for transcendental idealism can only be appreciated with reference to its emphasis on the original act and the orientation towards self-determination. The agency of reason, which involves the participation of various capacities, is an idealist antidote to the challenge of naturalistic scepticism, and it suggests a non-introspective notion of self-awareness.
4. Semantic externalism is incompatible with epistemic transparency. The incompatibility would make a strong case against transcendental idealism if the latter fails to distinguish between the apriority of self-knowledge and its infallibility. The apriority of self-knowledge is a result of the self-determining activity of reason, for which mental content *ought* to be transparent. The apriority does not warrant infallibility, and therefore transparency needs not *actually* obtain.

Reference

- Bell, David. 2003. "Transcendental Arguments and Non-Naturalistic Anti-Realism." In *Transcendental Arguments: Problems and Prospects*, edited by Robert Stern, 189–210. New York: Oxford University Press.
- Bilgrami, Akeel. 1992. "Can Externalism Be Reconciled with Self-Knowledge?" *Philosophical Topics* 20 (1): 233–67.
- Boghossian, Paul A. 1994. "The Transparency of Mental Content." *Philosophical Perspectives* 8: 33–50.
- Burge, Tyler. 1988. "Individualism and Self-Knowledge." *The Journal of Philosophy* 85 (11): 649–63.
- Burge, Tyler. 1996. "Our Entitlement to Self-Knowledge: I." *Proceedings of the Aristotelian Society New Series*, 96: 91–116.

- Cassam, Quassim. 1987. "Transcendental Arguments, Transcendental Synthesis and Transcendental Idealism." *Philosophical Quarterly* 37: 355–78.
- Cassam, Quassim. 2003. "Self-Directed Transcendental Arguments," In *Transcendental Arguments: Problems and Prospects*, edited by Robert Stern, 83–110. New York: Oxford University Press.
- Fichte, Johann Gottlieb. 1964–2012. *J. G. Fichte-Gesamtausgabe der Bayerischen Akademie der Wissenschaften*, edited by R. Lauth, H. Jacob, and H. Gliwitzky. Stuttgart-Bad Cannstatt: Friedrich Frommann Verlag.
- Fichte, Johann Gottlieb. 1982. *The Science of Knowledge*, trans. Peter Heath, and John Lachs. Cambridge: Cambridge University Press.
- Fichte, Johann Gottlieb. 1992. *Fichte: Foundations of Transcendental Philosophy: (Wissenschaftslehre) Nova Methodo (1796/99)*, trans. Daniel Breazeale. Ithaca, NY: Cornell University Press.
- Franks, Paul. 2003. "Transcendental Arguments, Reason, and Scepticism: Contemporary Debates and the Origins of Post-Kantianism." In *Transcendental Arguments: Problems and Prospects*, edited by Robert Stern, 111–46. New York: Oxford University Press.
- Harrison, Ross. 1982. "Transcendental Arguments and Idealism." *Royal Institute of Philosophy Lecture Series* 13: 211–24.
- Harrison, Ross. 1989. "Atemporal Necessities of Thought: or, How Not to Bury Philosophy by History." In *Reading Kant: New Perspectives on Transcendental Arguments and Critical Philosophy*, edited by Eva Schaper, and Wilhelm Vossenkuhl, 43–54. Oxford: Blackwell.
- Kant, Immanuel. 1998. *Critique of Pure Reason*, trans. P. Guyer, and A. Wood. Cambridge, UK: Cambridge University Press.
- Parent, Ted. "Externalism and Self-Knowledge." *The Stanford Encyclopedia of Philosophy*, Fall 2017 ed., edited by Edward N. Zalta. <https://plato.stanford.edu/archives/fall2017/entries/self-knowledge-externalism/>.
- Putnam, Hilary. 1975. "The Meaning of Meaning," In *Mind, Language and Reality; Philosophical Papers*, Vol. 2. Cambridge: Cambridge University Press.
- Schulze, G. E. 1985/1792. "'Aenesidemus', partial trans. by George di Giovanni." In *Between Kant and Hegel: Texts in the Development of Post-Kantian Idealism*, trans. and edited by George di Giovanni, and H. S. Harris, 105–35. Albany, NY: State University of New York Press.
- Stern, Robert. 2000. *Transcendental Arguments and Scepticism: Answering the Question of Justification*, New York: Oxford University Press.
- Strawson, P. F. 2007. *The Bounds of Sense*, New York: Routledge.
- Tse, Plato. 2018. "Fichte's Critique of Physicalism: Towards an Idealist Alternative." *Inquiry*, 62 (5): 527–45.
- Walker, Ralph C. S. 1978. *Kant*, London: Routledge & Kegan Paul.
- Zöllner, Günter. 2007. "From Transcendental Philosophy to *Wissenschaftslehre*: Fichte's Modification of Kant's Idealism." *European Journal of Philosophy*, 15 (2): 249–69.