

Edited by Halla Kim & Steven Hoeltzel

Transcendental Inquiry

Its History, Methods and Critiques



Transcendental Inquiry

Halla Kim • Steven Hoeltzel
Editors

Transcendental Inquiry

Its History, Methods and Critiques

palgrave
macmillan

Editors

Halla Kim
Department of Philosophy and Religion
University of Nebraska at Omaha
Omaha, Nebraska, USA

Steven Hoeltzel
Department of Philosophy and Religion
James Madison University
Harrisonburg, Virginia, USA

ISBN 978-3-319-40714-2
DOI 10.1007/978-3-319-40715-9

ISBN 978-3-319-40715-9 (eBook)

Library of Congress Control Number: 2016949045

© The Editor(s) (if applicable) and The Author(s) 2016

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Cover illustration: © EyeEm / Alamy Stock Photo

Printed on acid-free paper

This Palgrave Macmillan imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

1 Kant on the “Conditions of the Possibility” of Experience	1
<i>Claude Piché</i>	
2 Plato and Kantian Transcendental Constructivism	21
<i>Tom Rockmore</i>	
3 Kant and Fichte on the Notion of (Transcendental) Freedom	35
<i>Violetta L. Waibel</i>	
4 Fichte, Transcendental Ontology, and the Ethics of Belief	55
<i>Steven Hoeltzel</i>	
5 Transcendental Philosophy as “Therapy of the Mind”	83
<i>Benjamin D. Crowe</i>	
6 From Transcendental Philosophy to Hegel’s Developmental Method	107
<i>William F. Bristow</i>	
7 How Transcendental Is Cohen’s Critical Idealism?	127
<i>Halla Kim</i>	

8	Heidegger's Failure to Overcome Transcendental Philosophy	159
	<i>Eric S. Nelson</i>	
9	Others as the Ground of our Existence	181
	<i>Robert Stern</i>	
10	Raising Validity Claims for Reasons	209
	<i>Matthias Kettner</i>	
11	Transcendental Arguments Based on Question–Answer Contradictions	233
	<i>Yukio Irie</i>	
12	Consequences of the Transcendental-Pragmatic Consensus Theory of Truth	263
	<i>Michihito Yoshime</i>	
13	On Jürgen Habermas's Cognitive Theory of Morality	285
	<i>Yasuyuki Funaba</i>	
	Index	299

List of Abbreviations

Primary texts referenced repeatedly throughout the volume are cited in the notes using the following abbreviations.

Kant

- A/B *Kritik der reinen Vernunft* [Critique of Pure Reason], in AK, vols 3 and 4, trans. and ed. Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press, 1998).
- AK *Kants gesammelte Schriften* [*The Collected Works of Immanuel Kant*], 29 vols, ed. Royal Prussian (later German and Berlin-Brandenburg) Academy of Sciences, (Berlin: George Reimer (later Walter de Gruyter), 1900–).
- CPrR *Critique of Practical Reason*, in *Practical Philosophy*, trans. and ed. Mary J. Gregor (Cambridge: Cambridge University Press, 1996), 133–271.
- G *Grundlegung zur Metaphysik der Sitten* [Groundwork of the Metaphysics of Morals], in AK, vol. 4: 385–463.
- GMM *Groundwork of the Metaphysics of Morals*, in *Practical Philosophy*, trans. and ed. Mary J. Gregor (Cambridge: Cambridge University Press, 1996), 37–108.

viii List of Abbreviations

- P *Prolegomena zu einer jeden künftigen Metaphysik* [Prolegomena to Any Future Metaphysics], in AK, vol. 4: 253–383.
- PFM *Prolegomena to Any Future Metaphysics*, trans. Gary Hatfield, in *Theoretical Philosophy after 1781*, ed. Henry Allison and Peter Heath (Cambridge: Cambridge University Press, 2002), 49–169.

Fichte

- AD *J. G. Fichte and the Atheism Dispute (1798–1800)*, trans. Curtis Bowman, ed. Yolanda Estes (Burlington, VT: Ashgate, 2010).
- EPW *Early Philosophical Writings*, trans. and ed. Daniel Breazeale (Ithaca, NY: Cornell University Press, 1988).
- FNR *Foundations of Natural Right*, trans. Michael Baur, ed. Frederick Neuhouser (Cambridge: Cambridge University Press, 2000).
- GA *J. G. Fichte: Gesamtausgabe der Bayerischen Akademie der Wissenschaften* [The Complete Edition of the Bavarian Academy of the Sciences], 42 vols in 4 series, ed. Erich Fuchs, Hans Gliwitzky, Reinhard Lauth, Peter K. Schneider, and Günter Zöllner (Stuttgart-Bad Cannstatt: Frommann-Holzboog, 1964–2012).
- IWL *Introductions to the Wissenschaftslehre and Other Writings*, trans. and ed. Daniel Breazeale (Indianapolis: Hackett, 1994).
- SE *The System of Ethics*, trans. and ed. Daniel Breazeale and Günter Zöllner (Cambridge: Cambridge University Press, 2005).
- SK *Science of Knowledge, with the First and Second Introductions*, trans. and ed. Peter Heath and John Lachs (Cambridge: Cambridge University Press, 1982).
- StA *Die späten wissenschaftlichen Vorlesungen [The Later Academic Lectures]*, 3 vols, ed. Hans Georg von Manz, Erich Fuchs, Reinhard Lauth and Ives Radrizzani (Stuttgart-Bad Cannstatt: Frommann-Holzboog, 2000–).
- SW *Johann Gottlieb Fichtes sämtliche Werke [Johann Gottlieb Fichte's Collected Works]*, 8 vols, ed. I. H. Fichte (Berlin: Veit, 1845–46).
- VM *The Vocation of Man*, trans. Peter Preuss (Indianapolis: Hackett, 1987).

Hegel

- GW *Glauben und Wissen [Faith and Knowledge]*, in *Jenaer Schriften [Jena Writings] 1801–1807*, ed. Eva Moldenhauer and Karl Markus Michels, in vol. 2 of *Werke in zwanzig Bänden* (Frankfurt: Suhrkamp, 1970).
- PG Hegel, *Phänomenologie des Geistes [Phenomenology of Spirit]*, ed. Eva Moldenhauer and Karl Markus Michels, vol. 3 in *Werke in zwanzig Bänden* (Frankfurt: Suhrkamp, 1970).
- PS *Phenomenology of Spirit*, trans. A. V. Miller (Oxford: Oxford University Press, 1997).

Cohen

- ERW *System der Philosophie, Zweiter Teil: Ethik der reinen Willens [System of Philosophy, Second Part: Ethics of Pure Will]* (Berlin: Bruno Cassirer, 1904).
- KBE *Kants Begründung der Ethik [Kant's Foundations of Ethics]* (Berlin: Dümmmler, 1877).
- KTE *Kants Theorie der Erfahrung [Kant's Theory of Experience]* (Berlin: Dümmmler, 1871, 1885, 1918).
- LRE *System der Philosophie, Erster Teil: Logik der reinen Erkenntnis [System of Philosophy, First Part: Logic of Pure Knowledge]* (Berlin: Bruno Cassirer, 1902).
- PIG *Das Prinzip der Infinitesimal-Methode und seine Geschichte: Ein Kapitel zur Grundlegung der Erkenntniskritik [The Principle of the Infinitesimal Method and its History: A Chapter on the Groundwork of the Critique of Knowledge]* (Berlin: Dümmmler, 1883).
- W *Werke [Works]*, ed. H. Holzhey and H. Weidebach (Hildesheim: Olms, 2005).

Heidegger

- GA *Gesamtausgabe [Complete Edition]* (Frankfurt: Klostermann, 1976–).
- SZ *Sein und Zeit [Being and Time]*, 16th edn (Tübingen: Niemeyer Verlag, 1985). / *Being and Time*, trans. Joan Stambaugh (Albany: State University of New York Press, 1996).

Løgstrup

- ED *The Ethical Demand*, rev. trans. Hans Fink and Alasdair MacIntyre (Notre Dame: Notre Dame University Press, 1997).
- M1 *Metaphysics*, vol. 1, trans. Russell E. Dees (Milwaukee: Marquette University Press, 1995).

Habermas

- EA *Die Einbeziehung des Anderen* [*The Inclusion of the Other*], 2 vols (Frankfurt: Suhrkamp, 1997).
- MkH *Moralbewusstsein und kommunikatives Handeln* [*Moral Consciousness and Communicative Action*] (Frankfurt: Suhrkamp, 1983).

Introduction: What Hath Kant Wrought?

In his second-edition Introduction to the *Critique of Pure Reason*, Kant remarks, “I call all cognition transcendental that is occupied not so much with objects but rather with our mode of cognition of objects insofar as this is to be possible a priori. A *system* of such concepts would be called *transcendental philosophy*.”¹ These words mark a turning point in the history of Western philosophy, and today, more than two centuries later, intense interest abides in exploring and assessing transcendental approaches to philosophical problems—if not always as Kant himself would have envisioned or approved. On the systematic side, recent decades have seen a resurgence of rigorous work on transcendental arguments, understood as a special style of argumentation noteworthy for its antiskeptical ambitions,² as well as renewed interest in a broadly transcendental approach to issues in ontology.³ On the historical side, we have witnessed a remarkable revival of sympathetic

¹ AK 3:B25. English translation: *Critique of Pure Reason*, trans. and ed. Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press, 1998).

² See, for example, Robert Stern (ed.), *Transcendental Arguments: Problems and Prospects* (Oxford: Oxford University Press, 1999). Cf. Robert Stern, *Transcendental Arguments and Scepticism: Answering the Question of Justification* (Oxford: Oxford University Press, 2000).

³ See, for example, Markus Gabriel, *Transcendental Ontology: Essays in German Idealism* (London: Bloomsbury, 2011).

scholarly interest in post-Kantian idealism, including unprecedented attention to the transcendental philosophy of Fichte. And needless to say, Kant scholarship itself remains securely established as a sophisticated and impactful philosophical field in its own right. Moreover, the list of “transcendental” topics could easily be extended, as a glimpse at this volume’s table of contents will make plain.

The authors assembled here all undertake to enrich our systematic grasp and historical understanding of transcendental philosophy. Their chapters examine, among other things, the interestingly different forms that transcendental argumentation can take, the manifold uses to which transcendental methods can be adapted, and a wide array of provocative results reached along transcendental lines. Although the editors’ own views on such matters may occasionally show through in this overview, it has not been our aim, in compiling this collection or in authoring this introduction, to codify or lobby for any one understanding of transcendental arguments, transcendental methods, or transcendental philosophy “properly so called.” Our goal, instead, has been to bring a wide range of distinct outlooks and approaches into conversation, and thereby to document further the methodological and topical depth and breadth of transcendental inquiry.

Any thorough treatment of this subject should of course give careful consideration to Kant, and the present collection is no exception. In [Chap. 1](#), “Kant on the ‘Conditions of the Possibility’ of Experience,” Claude Piché examines a problem that has long puzzled Kant’s more incisive interpreters. Schematically stated, the issue is that the restrictions imposed on human knowledge by Kant’s transcendental epistemology seem to be violated by various assertions integral to that very theory. Claims that assign basic structural parameters to all possible experience, for example, or that posit noumenal causes for our sensory states, seem to deploy cognition’s pure categories (possibility, causality, etc.) for purposes of nonanalytic judgments concerning topics that transcend experience. Yet Kant’s official position on the categories’ legitimate theoretical employment seems to rule out any such use of these concepts. Thus it may appear that he cannot seriously state his own critical epistemology without implicitly (and uncritically) claiming nonanalytic insight into states of affairs that lie beyond the bounds of experience.

Piché argues, on the contrary, that Kant marks out the limits of valid a priori cognition in a way that does not inadvertently overstep those bounds. The argument hinges, in part, on the fact that, for Kant, our awareness of things' existence reflects our contingent, receptive relationship to a sensory manifold whose contents do not suffice for truth-aptness cognition of extrasubjective objects or relationships—"experience," in Kant's thick sense of that term. Accordingly, the a priori conditions for the possibility of experience are themselves conditioned, and in more than one way: their deployment is contingent upon sensory manifestations whose occurrence and content are not under our intellectual control, and their pure content is a function of the range of ways in which sensory intake underdetermines even the most elementary empirical cognition. Kant's claims about experience, its possibility, and its a priori conditions are thus all anchored in a more basic acknowledgment of human intellection's inescapable limitations. As such, these claims do not presuppose or pretend to any power of unconditioned insight.

Kant's transcendental philosophy revolves around the claim that "the conditions of the *possibility of experience* are at the same time conditions of the *possibility of the objects of experience*."⁴ Evidently, then, for Kant transcendental cognition essentially involves reflection on the nexus that connects (first-order, nontranscendental) acts of cognition with the objects to which these acts refer. The basic issue here is a vexed and a venerable one—take for example Parmenides's momentous assertion of the identity of knowing and being—and scholars have yet to reach consensus about the content of Kant's position on this question. Witness for instance the continued dispute over whether Kant defends (i) a "two-objects" theory, according to which the objects of empirical cognition are mind-dependent appearances, as opposed to mind-independent (and wholly unknowable) things in themselves; or instead (ii) a "two-aspects" account, according to which the terms "appearances" and "things in themselves" designate, not two distinct classes of objects, but two radically different types of descriptions under which any one object can be taken.

⁴ A158/B197.

In Chap. 2, “Plato and Kantian Transcendental Constructivism,” Tom Rockmore argues that we can make headway here by considering the ways in which Kant reacts to and corrects a Platonic position on the nature and powers of cognition. Rockmore does not depict Kant as an orthodox Platonist, or Plato as some sort of proto-Kantian; instead, he underscores Kant’s deep debt to Plato’s prior rejection of the inference from sensible appearances to a mind-independent reality which—according to the representational realist, but not for Plato or for Kant—causes and resembles those appearances. Kant of course rejects Plato’s further claim that the intellect can intuit ultimate reality. But he does not, Rockmore argues, settle for skepticism in consequence of our inability to infer or intuit “the real as such.” Instead, he develops a transcendental account of cognition’s constructive relation to “the real for us,” which still allows for modest claims to knowledge.

The above discussions principally concern the epistemological side of the Kantian project. But Kant is also a towering figure in practical philosophy, and his outlook on this topic centers on a complex account of the nature and possibility of “transcendental freedom.” In contrast with a broadly compatibilist, merely “psychological and comparative” kind of freedom (which for Kant does not suffice to make actions morally imputable),⁵ this “*transcendental freedom* . . . must be thought as independence from everything empirical and so from nature generally.”⁶ Accordingly, a will that possesses such independence must be one that can determine itself to action based on purely rational considerations that have substantive moral implications.⁷

These ideas will of course be familiar to many readers. It may be less well known that in the German-speaking philosophical world of the early 1800s, it was generally held that the distinctly Kantian approach to ethics was best exemplified not by Kant’s work but by Fichte’s.⁸ In

⁵ See Kant, *Critique of Practical Reason*, AK 5:95–8; CPrR 216–18.

⁶ AK 5:97; CPrR 217.

⁷ See, for example, AK 5:28–30; CPrR 162–4.

⁸ See Michelle Kosch, “Fichtean Kantianism in Nineteenth-Century Ethics,” *Journal of the History of Philosophy* 53, no. 1 (2015): 111–32.

Chap. 3, “Kant and Fichte on the Notion of (Transcendental) Freedom,” Violetta L. Waibel examines Fichte’s analysis of freedom into three main stages, and then charts some relationships between this account and Kant’s earlier analysis of freedom into two main types. Consciously building on Kant’s position, Fichte distinguishes between formal, material, and moral stages of freedom, understood as increasingly adequate articulations of reason’s basic spontaneity and the rational being’s basic commitment thereto. Formal freedom flows from our capacity for conceptual thinking, whereby we can place ourselves at a reflective remove from our merely given impulses and drives. For Fichte, of course, all thinking involves an element of willing, and vice versa, so this reflective distancing always has an evaluative dimension, and in material freedom, the next stage, this dimension matures into an active principle of motivation and orientation, as the subject comes to recognize better and increasingly prioritize its own capacity for independence. Finally, in moral freedom, the aim of attaining independence gains additional, specifically ethical, focus through the subject’s rational recognition that truly unmitigated autonomy can be achieved only via principled moral agency. Waibel argues that the diverse aspects of freedom brought to light by Fichte’s three-stage analysis, while not all explicitly upheld as such by Kant, are all implicitly ingredients in the latter’s account of freedom as the autonomy of reason.

Fichte’s *Wissenschaftslehre* will raise in any careful reader’s mind many interesting questions about the possibilities and boundaries of transcendental philosophy. As an especially contentious case in point, what ought we to make of Fichte’s seeming pivot, somewhere around 1800, to an overtly metaphysical position? Does the broadly theistic immaterialism boldly proclaimed in *The Vocation of Man* signal a failure to keep faith with the Kantian project? Or is that merely a misleading presentation, in ontologically lofty language, of a position that remains, in principle, ontologically noncommittal? Or is there some way in which to defend such metaphysical commitments on transcendently respectable grounds? In Chap. 4, “Fichte, Transcendental Ontology, and the Ethics of Belief,” Steven Hoeltzel offers an argument for that third alternative.

Fichte's epistemology, on this account, employs transcendental arguments of an essentially Kantian kind—regressive arguments that isolate necessary conditions for the possibility of experience—in order to secure two key conclusions: (i) that metaphysical issues concerning the existence and nature of extrasubjective reality are epistemically intractable; and (ii) that pure reason is practical, and in a way that gives us nonepistemic but still quintessentially rational grounds for assent to a specific metaphysical outlook. The latter will be an epistemically indefeasible outlook describing a supersensible order of being: one in which approximation to pure reason's highest ideal is actually possible, such that the affiliated rational requirements retain real prescriptive power. On Hoeltzel's reading, then, Fichte's transcendental epistemology issues in a distinctive conception of the basic nature of rationality and rational justification, and the latter supports a distinctly Fichtean form of pure-rational nonevidentialism, in the light of which a steadfastly transcendental philosophy can justifiably comprise some firm convictions concerning certain basic metaphysical questions. (Nonevidentialism is the view that in certain cases we are required to assent to propositions for which our evidence is insufficient, and on this transcendentially founded Fichtean form of nonevidentialism, pure reason is the source of the said requirement.)

Another way in which Fichte's philosophy may prompt reflection on the nature and scope of transcendental inquiry is in its explicit identification of specifically transcendental philosophy with a decidedly non-standard viewpoint. This is a cognitive stance radically removed from the unreflective outlook to which we default in everyday life. Moreover, as a philosophical stance it is nothing like the naturalism that accepts the essentials of our prephilosophical point of view, seeking only to add nonempirically to the knowledge that the said outlook supposedly affords. On the contrary, Fichte proposes to suspend transcendentially, and then to reconceive philosophically and supersede, our prephilosophical outlook as such and as a whole.

Benjamin D. Crowe pursues this theme through some of Fichte's lesser-known later writings in [Chap. 5](#), "Transcendental Philosophy as 'Therapy of the Mind': Fichte's 'Facts of Consciousness' Lectures." Crowe's examination focuses on two of Fichte's main inspirations: the

Kantian conception of reason as essentially epigenetic, and the ancient conception of philosophy as importantly therapeutic. On Crowe's account, Fichte's transcendental *medicina mentis* aims to help us to raise ourselves above our everyday outlook and, eventually, to grasp our own (and every other) consciousness's belonging to a single, epigenetically unfolding pure life. Fichte conducts this progression, Crowe argues, via a series of quasi-transcendental arguments which, starting from some already-manifest structure or fact, propose to disclose some deeper structure within which the former is embedded and by which it is qualified. So, for example, the prephilosophical self-conception of empirical consciousness as an inert, passive mirroring should give way to a (transcendental) understanding of such consciousness as a dynamic, living imaging—which imaging, in turn, must be understood to implement a prior purposive orientation, which then entails the thought of a unifying moral nexus, and so on, until one finally arrives at the philosophical vision that beholds, in the totality of finite becomings, the manifestation of one free, divine life.

In Chap. 6, “From Transcendental Philosophy to Hegel's Developmental Method,” against the widespread inclination to interpret Hegel's method as a modified version of Kant's, William F. Bristow highlights the opposition between Kant's transcendental method and Hegel's method in the *Phenomenology of Spirit*. He argues that Hegel rejects Kant's method, principally on the grounds that Kant's method presupposes the subjectivism—the view that we can know things only as they are for us, not as they are in themselves—characteristic of his transcendental idealism. Hegel also cannot accept the dualism and formalism implied by Kant's view. Bristow goes on to show how Hegel configures his own method in the *Phenomenology* to be free of the sort of fixed starting point from which Kant's transcendental arguments proceed. Hegel's developmental method is one of the most intriguing aspects of his philosophy, but it can only be appreciated to the extent that the temptation is resisted to interpret the method according to a Kantian paradigm. While Kant's method proceeds from some common ground, something fixed and unquestioned among disputants, Hegel's method is designed to function in the situation in which there is no recognized common ground. Thus it begins with a

context in which nothing is taken as fixed or established. This idea is at the foundation of what Bristow calls “Hegel’s criterion-less critique.”

Further, for Hegel, through the progress of human knowledge in history, human subjectivity undergoes distinctive transformations. And in this dialectical movement, new forms of objectivity emerge for consciousness in what Hegel calls “experience.” Going beyond the Kantian standpoint that views the finite, criticizing subject as fixed, Hegel’s developmental method does not take anything for granted as a fixed point from the outset. Bristow suggests that this approach has an advantage over Kant’s because “it enables philosophical critique to proceed without prior commitments, in maximal openness regarding the outcome.” Thus Hegel’s method achieves an outcome (“the resting place,” as Bristow puts it) only through describing a self-reverting circle, encompassing all significant positions as stages in spirit’s self-knowing and self-becoming. Even though this chapter limits itself to the *Phenomenology*, it has interesting repercussions and can shed light on the arguments and positions Hegel advances in his *Science of Logic*.

We move on to examine some later developments in transcendental philosophy in Chap. 7, “How Transcendental Is Cohen’s Critical Idealism?,” by Halla Kim. Neo-Kantianism, long belittled as an out-of-date and obsolete academic movement, is slowly making a comeback. In this chapter, Kim discusses the critical idealism of Hermann Cohen (1842–1918), one of its founding members, and examines the extent to which it owes to, but also goes beyond, Kant’s transcendental philosophy. Kim suggests that, while remaining within the venerable tradition of transcendental philosophy, Cohen goes well beyond Kant and reveals a new dimension of that philosophy.

Kim starts by introducing Kant’s and Cohen’s conceptions of the transcendental. He explains how Cohen’s appropriation of Kant’s transcendental philosophy, especially in his early *Kant’s Theory of Experience*, leads to a distinctive transcendental method which adapts the analytic method employed in Kant’s *Prolegomena*. This is how Cohen derives the necessary a priori conditions for the objectivity of mathematical-natural knowledge to his satisfaction. It is then shown that this method, while capable of isolating the underlying conditions of experience, cannot be fully ground-laying. But for the purpose of

ultimate grounding, Cohen also cannot accept Kant's synthetic method (mainly employed in the *Critique of Pure Reason*) because of its psychologism. Kim suggests that Cohen completes the transcendental method in a new way by means of what might be called "the method of hypothesis" and the ensuing doctrine of origin. Cohen here makes a transition from Kant's transcendental idealism to his own critical idealism by way of his recourse to Plato. In particular, for Cohen, a hypothesis (a concept that Cohen recovers from Plato) is a thesis that presents the point of logical commencement for a proper philosophical procedure. Cohen then identifies a hypothesis with what he terms "origin" (*Ursprung*). The principle of origin then must be ground-laying (*grundlegend*). As the ground-laying principle, origin must be presuppositionless, unconditioned, and necessary—"a systematic point of culmination" in Cohen's systematic idealism. But this origin is not a product, nor a stationary entity or its state. It is an activity, a movement, and a creative vitality in thinking. Kim argues that, in abandoning the synthetic method in favor of the method of hypothesis within the framework of Kant's analytic method, Cohen's critical idealism goes well beyond the purview of the Kantian project with his recourse to origin.

Neo-Kantianism is of course not the only latter-day attempt to overhaul and update the transcendental project inaugurated by Kant. We also have transcendental phenomenology, for example, most notably as pursued in the pivotal work of Husserl. And yet Heidegger, who may well be the most momentous single figure in philosophy since Kant, launches his own project largely by rejecting transcendental phenomenology in the Husserlian mode. In doing so, does Heidegger seek to break free from transcendental inquiry in general? And supposing that this is his objective, does he actually accomplish it?

These prove to be complicated questions, and in [Chap. 8](#), "Heidegger's Failure to Overcome Transcendental Philosophy," Eric S. Nelson explores Heidegger's ambivalent and ambiguous relationship with that philosophy. Clearly, Heidegger sought from the start to overcome what he saw as the overly static and ahistorical accounts of the constitution of meaning characteristic of the prior transcendental tradition. And the so-called "turning" of the mid-1930s, in which he

explicitly rejects his own earlier overemphasis (in and around *Being and Time*) on the meaning-constituting role of the subjectivity of the subject, only seems to distance Heidegger further from any transcendental standpoint. Nelson argues, however, that after the turning Heidegger continually alternates between, on the one hand, a rhetoric of radically overcoming transcendental philosophy along the above lines (and thereby twisting free of the more problematic tendencies of Western metaphysics and modernity) and, on the other hand, a reconception of transcendental constitution in terms of the event (*Ereignis*) of being, world, and history—thus remaining within the transcendental tradition even while abjuring some of its earlier articulations. All things considered, Nelson suggests, transcendental philosophy may still provide us with the best lens through which to view the whole of Heidegger’s work, despite the latter’s own anti-transcendental self-interpretations.

Heidegger’s thought has for various reasons provoked very strong reactions. Interestingly, some of the most impactful of them can be formulated in transcendental terms. In [Chap. 9](#), “Others as the Ground of Our Existence: Levinas, Løgstrup, and Transcendental Arguments in Ethics,” Robert Stern examines a transcendental reading of some central dimensions of Levinas’s ethics,⁹ contrasting this approach with Løgstrup’s endeavor to secure some kindred conclusions in a less ambitious—but still markedly transcendental, and arguably more successful—way. According to the relevant reading of Levinas, his goal is to refute moral skepticism by showing that our own identities and values presuppose a profound prior indebtedness to others. Consequently, we cannot doubt or deny others’ deep importance for us without tacitly drawing upon our inescapable debt to them, thereby performatively contradicting our moral-skeptical claims. However, Stern argues that such reasoning does not really suffice to silence the skeptic, because it leaves open the question why an ethical commitment that one has to have already enacted on some prereflective level ought also to be regarded as a commitment that one is validly required to further endorse from the standpoint of reflection.

⁹ Diane Perpich, *The Ethics of Emmanuel Levinas* (Stanford: Stanford University Press, 2008).

Outlining an alternative approach, Stern suggests that Løgstrup (and perhaps also Levinas) points up our deep debts to others, not in order to respond directly to the skeptic, but in order to expose and critique a type of basic oversight which one has to commit before one can really take moral skepticism seriously. The goal of this approach is to show that simply to regard moral skepticism as a serious candidate for acceptance—an outlook that may be shared both by the skeptic *and* by those whose propose to meet the skeptic’s challenge head-on—is already to have overlooked some deep fact about ourselves which transcendental reflection can help bring to light, and in the light of which the skeptic’s questions appear egregiously ill-considered. For Løgstrup, this is made plain via transcendental considerations concerning the ways in which any party to human life as we know it is always already deeply dependent upon relations of trust, communication, and care, and responsive to the norms that structure and sustain such relations.

Matthias Kettner, a champion of Karl-Otto Apel’s transcendental pragmatics of communication, offers a thorough and rigorous philosophical foundation for a discourse-ethical approach to meta- and normative ethics in [Chap. 10](#), “Raising Validity Claims for Reasons: Transcendental Reflection in Apel’s Argumentative Discourse.” The goal is to establish that the dialogical practice of fully engaged argumentative discourse necessarily involves conceptually normative presuppositions, some of which have a universally valid and recognizably moral content. For this purpose, Kettner works to identify conceptually normative presuppositions of argumentation, select those that are morally charged, and then develop their moral content into a coherent core conception of a morality with unassailable rational credentials. The upshot, if successful, is that, via reflexive recourse to practices of discursive argumentation, we can ground in a rationally definitive way (*letztbegründend*) certain normative requirements (moral and other) which rational persons as such must meet. In the process, Kettner, following Apel (but unlike Habermas), elaborates from the notion of a performative self-contradiction a method of rationally definitive justification.

Can this project succeed? Evidently it all depends on how successfully one can show that validity is not a metaphysically isolated

phenomenon but something that emerges from communicatively constructed claims for the purpose of disclosing those normative commitments, roles, and relations in the reciprocally shared self-understanding of discourse-participants that are indispensable for the potentially rational powers of argumentative discourse. Suggesting that the idea of claiming validity intrinsically involves the idea of a social practice, Kettner characterizes argumentative discourse essentially in the following way: “Argumentative discourse is a social practice, open to all persons in their capacity as reasonable evaluators, which has as its aim the communicatively rational revision of conflicting reasons with (apparently less conflicting) reasons.”

Note that this characterization operates with reasons as such, and in this respect improves on the more conventional conception of reasons as justifiers of the three or four “universal validity claims” such as truth, rightness, veracity, and meaning, to which Apel and Habermas repeatedly refer: it has more power and versatility in being more general and having wider scope. Kettner concludes with some remarks on the prospects of discourse ethics, according to which the only unassailable hope for true moral progress is hope in the progressive globalization of the ethos of discourse.

In [Chap. 11](#), “Transcendental Arguments Based on Question–Answer Contradictions,” Yukio Irie explores an alternative approach to transcendental grounding in discourse ethics. For Irie, a substantial part of language is best understood as based on relations of mutual linguistic responsiveness among its practitioners. What are the fundamental conditions for these relations? Irie suggests that perhaps the best method for addressing such problems relies on a particular linguistic phenomenon that appears to reveal some fundamental conditions underlying linguistic communication. This phenomenon involves a contradiction based on a question and an answer (QA contradiction, for short), which can be illustrated as follows: “Can you hear me?”—“No, I cannot hear you.” The QA contradiction seems to reveal fundamental structures or conditions of communication, because a substantial part of linguistic communication is essentially constructed out of relations of questions and answers and their ilk; hence, the contradiction points to some basic or transcendental conditions for communication. Through a series of

careful analyses of QA contradictions, Irie points out that unlike most transcendental arguments based on classical logic, his analysis reveals a different type of transcendental argument, one which has important repercussions for our understanding of the normativity of language and the principles of identity, mutual belief, and mutual recognition, *inter alia*.

Irie suggests that the method he proposes is a transcendental argument based on QA contradictions whose transcendental conditions are more fundamental than the conventional semantic/cognitive presuppositions. For the former present universal conditions, whereas the latter are empirical and vary case by case. As to the nature of transcendental conditions, it might appear that we can justify our knowledge and ethical principles by reflecting on QA contradictions in the way that, for example, Apel tried to justify them by reflecting on pragmatic contradictions. Irie, however, finds such foundationalism dubitable, much like Rorty, who, while acknowledging the force of transcendental argument, admits it only as a parasitic argument which always presupposes some theory in order to prove the necessity of a conclusion. Irie is inclined to hold that transcendental argument represents only a weak form of justification.

Among the controversial fallout of the theory of transcendental pragmatics proposed by Apel is the so-called “consensus theory of truth,” which has been widely subject to criticism. This theory regards truth as the ultimate consensus of the ideal communication community, building upon the Peircean consensus theory and accepting Habermas’s latter-day variation. One can readily discern the influence of Peirce’s semiotic interpretation of Kant on Apel’s formulation of transcendental pragmatics. Peirce’s consensus theory of truth, which is combined with the Kantian interpretation, is another of the elements that determined the fundamental direction of Apel’s project.

In [Chap. 12](#), “Consequences of the Transcendental-Pragmatic Consensus Theory of Truth,” Michihito Yoshime considers two criticisms of the consensus theory of truth, one by Wellmer and another by Putnam. Yoshime suggests that there is limited validity to these criticisms; however, rather than directly responding to them, he leverages these criticisms for the purpose of clarifying the transcendental-pragmatic

notion of truth and illustrating its relevance to ultimate grounding, with the result that these criticisms become less compelling. In particular, he points out a metaphysical-realistic presupposition implicit in the criticisms, and then evaluates them by placing the transcendental-pragmatic consensus theory of truth in its proper, antirealistic framework, in opposition to that metaphysical-realistic outlook. Finally, Yoshime examines the meaning of “ultimate” grounding in light of a transcendental, antirealistic understanding of truth.

Habermas is well-known for defending moral cognitivism, the view that moral judgments have cognitive content analogous to truth value, despite his having no sympathy for moral realism. In [Chap. 13](#), “On Jürgen Habermas’s Cognitive Theory of Morality,” Yasuyuki Funaba examines Habermas’s view and traces it to the latter’s differentiation of assertoric sentences used in constative speech acts from normative sentences used in regulative speech acts. According to Habermas, the principle of universalization (U) is the moral principle that regulates moral argumentation, because it expresses the injunction to maintain the impartial moral point of view, constraining all participants to adopt the perspective of all others:

(U) All affected can accept the consequences and the side effects that [the norm's] general observance can be anticipated to have for the satisfaction of everyone's interests, and the consequences are preferred to those of known alternative possibilities for regulation.

This is the condition for the validity of all acceptable moral norms. No moral norms then can be grounded in an absolute way (*letztbegründet*) if one assumes a cognitivist theory; instead, one can only say that for some moral norms there are just no alternatives. And this lack of alternatives manifests itself again and again through all moral argumentation.

Funaba then goes on to consider two of Putnam’s criticisms of this view. The first is that arguments do not always end in moral consensus. Imagining a father who is cruel to his child, Putnam charges that one cannot reach a consensus on the right moral norms, even when following (U) and the rules of argumentation. Our moral sense and intuitions cannot be reduced to a set of rules for settling disputes. Funaba suggests

that Putnam here wants to see the rules of argumentation as sufficient conditions for the right moral norms. However, Habermas cannot accept this, because he demands that in all argumentation the validity of the rules of argumentation itself is questioned and validated again and again. The validity of (U) as the moral principle can always be justified at the same time as the validity of the presuppositions as rules is confirmed by the argumentation, by pointing out the performative contradiction committed by any arguer who objects to the validity of these rules. The second of Putnam's criticisms is that Habermas's theory is a moral minimalism that tries to reduce moral norms "strictly" to the rules of argumentation. Questioning the fact/value dichotomy, Putnam contends that Habermas, with regard to the case of the cruel father above, can only say, "Discuss according to the rules of argumentation the question whether you are allowed to be cruel or not!" However, he cannot say, "Don't be so cruel!" Funaba responds that it goes without saying that Habermas cannot accept this charge, because, as pointed out above, he differentiates between constative and normative statements.

Steven Hoeltzel
Halla Kim

1

Kant on the “Conditions of the Possibility” of Experience

Claude Piché

In what follows, I would like to bring out some features of Kant’s conception of transcendental philosophy in the *Critique of Pure Reason*. One of the aims of this work was to answer the question raised in the famous letter to Markus Herz of February 21, 1772, about the possibility of a priori cognition’s relating to its object.¹ And as we know, the answer elaborated in the *Critique* restricts the use of this pure cognition to possible experience. Cognition a priori is valid only when applied to appearances, within experience. It leads therefore to a finite knowledge, limited to the phenomenal world.

Now the question is: What is the legitimacy of Kant’s own transcendental discourse when we take account of the restrictions imposed on

¹Letter to Markus Herz, AK 10:131. Kant focuses his question here on the “intellectual representations.”

C. Piché (✉)

Department of Philosophy, Université de Montréal, Montreal, Canada

human knowledge as a result of this very investigation? In other words, the use of a priori cognition being restricted to possible experience, how could the cognitive claim of this philosophical inquiry be justified if it does not itself take place within possible experience? The problem already becomes obvious with the formulation of the main result of the critical investigation, namely that the only valid sphere for a priori cognition is “possible experience.” Yet how are we to understand the use of the word “possible” here? To be sure, possibility is a modal category whose conditions of legitimate application are strictly determined and restricted in the *Transcendental Analytic*. And the same goes for the category of contingency, which, as we will see, plays an important role in specifying the status of the conditions of the possibility of experience. But then again the question is: Does Kant make an inappropriate or illegitimate use of these categories when, rather than applying them to objects of a possible experience, he uses them to describe the scope of valid a priori knowledge as a whole? By apparently removing the restrictions on their use here, it seems as though Kant is making a transcendent or, as he would say, “transcendental,” employment of the categories of modality.

I will argue, however, that this use of the categories complies, all things considered, with the constraints imposed upon them for their application in experience. We will find that the categories at work in Kant’s transcendental discourse are not employed without caution in their purely intellectual significance, which would lead them to open unto the unconditioned, as in the *Transcendental Dialectic*. On the contrary, the standpoint from which critical philosophy condemns every attempt to gain knowledge of the unconditioned is not itself unconditioned. I would like to show that transcendental philosophy, even when it circumscribes the limits of human knowledge with the help of the modal categories of possibility and contingency, nevertheless acknowledges its own finitude and refrains from overstepping the “bounds of sense.” In order to achieve this, we will have to attend clearly to and carefully define Kantian expressions found in the first *Critique*, such as “possible experience,” “transcendental cognition,” “conditions of the possibility,” and “contingency,” which in the end will lead us to establish the modal status of the ultimate presuppositions of critical philosophy.

Possible Experience

Before we inquire into the role played by possible experience in the transcendental philosophy developed in the first Critique, we would be well advised to focus on the meaning of each of these words. “Experience” is a well-known Kantian term designating the cognition of objects that are “given in empirical intuition” or, if one prefers, the “cognition of the objects through perception.” The short definition is thus “empirical cognition,”² whereby it is understood that this cognition lays claim to objectivity. Experience is in fact objective empirical knowledge.

“Possibility” is a modal category which, from a strictly logical point of view, refers to that which does not contradict itself. To say that a concept is “possible” simply means that its internal components are not in contradiction to each other. This is what Baumgarten in his *Metaphysica* calls “absolute” possibility³—a characterization for which he is criticized in Kant’s lectures on metaphysics as well as (though implicitly) in the Transcendental Dialectic.⁴ In fact, Baumgarten uses the adjective “absolute” to designate “intrinsic” possibility, which according to Kant is the “least” that can be said of the concept of an object. But the strong meaning of the word should be maintained. In Kant’s view, absolute possibility properly signifies the “most” that can be said of a concept, namely that its object is possible literally “in all respects” and “without any restriction” whatsoever.⁵

Kant feels the need to restore the strong sense of the term “absolute” at the beginning of the Transcendental Dialectic because it is synonymous with “unconditioned,” which is the main topic of

² B289, B219, and B165–6.

³ Alexander Gottlieb Baumgarten, *Metaphysica* (4th edition, 1757), §15, reproduced in AK 17:29.

⁴ Kant, *Vorlesungen über Metaphysik (Volckmann)*, 1784–85, AK 28:406; A324–6/B380–2; A232/B284.

⁵ A324–6/B381–2. In *Reflexion* 4297, Kant even goes so far as to claim that “*Was in aller Absicht möglich ist, ist wirklich*” (AK 17:499). See on this topic Burkhard Hafemann, “Logisches Quadrat und Modalbegriffe bei Kant,” *Kant-Studien* 93, no. 4 (2002): 415.

this second part of the *Critique* devoted to the logic of illusion. Indeed, a possibility that is absolute in the full sense of the word pertains not to the understanding but to reason,⁶ and this category gives rise to the same problem as the modal concept of “unconditioned necessity” used by the dogmatic metaphysician in the cosmological proof of the existence of God. That is, both concepts exceed the grasp of the human mind: “the unconditioned necessity, which we need so indispensably as the ultimate sustainer of all things, is for human reason the true abyss.”⁷ The treatment reserved for the modal categories in the Transcendental Analytic reveals that only “hypothetical” necessity, that is, conditioned necessity, is available to human cognition, and the same goes for possibility, as we learn in *Reflexion* 4005: “with reason we can cognize only conditioned possibility.”⁸ The only kind of possibility that can be grasped by a finite understanding is relative possibility, namely that which is possible only in some respects. Accordingly, possibility is cognizable only if it is “restricted to conditions.”⁹ Let us take for instance Kant’s example of the “invented concepts” of substances and forces supposedly present in experience. It is not enough to say that such concepts are possible because they are not self-contradictory. To be sure, this satisfies the minimal requirement of their logical possibility, but their real possibility also has to be established within experience by showing that these objects can be instantiated according to the known laws of experience.¹⁰ This is a clear example of conditioned possibility.

⁶ A232/B285.

⁷ A613/B641. See Toni Kannisto, “Modality and Metaphysics in Kant,” in *Kant und die Philosophie in weltbürgerlicher Absicht: Akten des XI. Kant-Kongresses 2010*, ed. Stefano Bacin, Alfredo Ferrarin, Claudio La Rocca, and Margit Ruffing (Berlin: de Gruyter, 2013), 639.

⁸ A228/B280; *Reflexion* 4005, AK 17:382.

⁹ A326/B382. “Die hypothetische Möglichkeit [ist], als eine kleinere Möglichkeit zu betrachten, weil sie immer nur unter Restriktion statt findet” Kant, *Vorlesungen über Metaphysik (v. Schön)*, 1780?, AK 28:488.

¹⁰ A222–3/B269–70.

Now when Kant considers possible experience as a whole, rather than just particular objects within it, he maintains a similar restriction on his recourse to the concept of possibility. Experience might well be declared “possible,” but here again only as a conditioned possibility. Hence the central expression for the *Critique of Pure Reason*: “conditions of the possibility of experience.”¹¹

¹¹ A158/B197, my emphasis. As their titles indicate, the following studies deal with the same topic as the present chapter, but they do not adopt the approach proposed here: A. R. Raggio, “Was heisst ‘Bedingungen der Möglichkeit?’,” *Kant-Studien* 60, no. 2 (1969): 153–65; Peter Struck, “Kants Formel von den Bedingungen der Möglichkeit von...und die Ableitung der transzendentalen Einheit des Selbstbewusstseins,” *Prima Philosophia* 6 (1993): 257–66; Arthur Collins, *Possible Experience: Understanding Kant’s “Critique of Pure Reason”* (Berkeley: University of California Press, 1999). However, in the collection of essays edited by Eva Schaper and Wilhelm Vossenkuhl under the title *Bedingungen der Möglichkeit: “Transcendental arguments” und transzendentes Denken* (Stuttgart: Klett-Cotta, 1984), there is an interesting contribution from Rüdiger Bubner on the self-referentiality of Kant’s transcendental argumentation entitled “Selbstbezüglichkeit als Struktur transzendentaler Argumente,” 63–79. This is in effect the perspective that I adopt here, although I do not develop it in the same way as Bubner does; he applies it to the different levels of “synthesis” in Kant’s transcendental deduction. For my part, I see self-referentiality in Kant’s applying *mutatis mutandis* the restrictions on the use of the categories in experience to his own transcendental argumentation. In any event, the general conception of self-referentiality laid out by Bubner in an earlier version of his thesis remains valid for my own undertaking:

According to Kant the cognitions that may be called transcendental are only those in which cognition is treated in relation to its specific possibilities. Consequently, the cognition that is called transcendental thematizes together with the universal conditions of cognition the presuppositions of its own emergence and functioning. Self-referentiality is characteristic of the transcendental argument. If one can show that the reasoning on factual forms of cognition and the explanation of their presuppositions is impossible without having recourse to certain elements of these very forms of cognition, then not only at the level of the factuality of cognition will a state of affairs be demonstrated, but at a higher level the constant validity of these universal forms of cognition will be confirmed. (Transzendental dürfen Kant zufolge nur Erkenntnisse heissen, in denen die Erkenntnis in bezug auf ihre spezifischen Möglichkeiten Thema ist. Wenn dies gilt, so thematisiert die transzendental genannte Erkenntnis mit den allgemeinen Erkenntnisbedingungen auch die Voraussetzungen ihres eigenen Entstehens und Arbeitens. Für das transzendente Argument ist die Selbstbezüglichkeit kennzeichnend...Wenn sich zeigt, dass das Raisonement über faktische Erkenntnisformen und die Aufklärung von deren Voraussetzungen ohne Benutzung gewisser Elemente jener Erkenntnisformen unmöglich ist, so wird nicht bloss auf der Ebene der Faktizität von Erkenntnis ein faktischer Umstand demonstriert, sondern auf einer Metaebene die ungebrochene Geltung allgemeiner Formen des Erkennens bestätigt.)

See his “Zur Struktur eines transzendentalen Arguments,” in *Akten des 4. Internationalen Kant-Kongresses*, Teil 1, ed. Gerhard Funke (Berlin: de Gruyter, 1974), 23, 25. I have also published on

Experience taken globally will thus be possible solely under a set of conditions. As we can see, Kant remains coherent in his use of the category of possibility: experience's possibility also depends on conditions, and most importantly on transcendental conditions. We will have to inquire into the nature of this transcendental conditioning, but before going any further we must again focus our attention on Kant's terminology.

Transcendental Cognition

Let us quote the canonical definition of the term "transcendental" given in the Introduction to the second edition of the *Critique*: "I call all cognition transcendental that is occupied not so much with objects but rather with our mode of cognition of objects insofar as this is to be possible a priori."¹² Here we clearly have two levels of cognition: a first level concerned with the cognition of objects and a second level dealing with our mode of cognition a priori of these objects, which is precisely what transcendental cognition is about. This means that in the Transcendental Analytic, Kant is concerned primarily with a priori cognition, or better: with our a priori mode of cognition of objects.

However, all a priori cognition is subject to a constraint: it can never reach the object in its actuality. For this, according to the teachings of the Transcendental Analytic, the empirical dimension of the object must be added, at which point the cognition becomes a posteriori. As we can read in the *Metaphysical Foundations of Natural Science*, "to cognize something a priori means to know it in its mere possibility."¹³ Transcendental cognition is no exception here, and so it is no surprise

this subject: "Self-Referentiality in Kant's Transcendental Philosophy," in *Proceedings of the 8th International Kant Congress*, Book 2.1, ed. H. Robinson (Milwaukee: Marquette University Press, 1995), 259–67; and "La dimension autoréférentielle du discours sur les 'conditions de possibilité,'" in *Kant*, ed. Jean-Marie Vaysse (Paris: Éditions du Cerf, 2008), 191–211.

¹² B25. See Tinca Prunea-Bretonnet, "De l'ontologie à la philosophie transcendantale: dans quelle mesure Kant est-il wolffien?," in *Kant et Wolff: Héritages et ruptures*, ed. Sophie Grapotte and Tinca Prunea-Bretonnet (Paris: Vrin, 2011), 160.

¹³ AK 4:470.

that it is oriented exclusively toward “possible” experience.¹⁴ The a priori cognition introduced in the Transcendental Analytic is aimed at grounding the mere possibility of experience, and for this purpose it has to “abstract from everything empirical in the appearances.”¹⁵

Furthermore, it must be noted that cognition of the a priori conditions of experience is not just any kind of a priori cognition. It is, as I have already said, a second-level cognition. Its distinctive nature will become clear if we compare it with the first-level kind of a priori cognition such as pure geometry. While geometry is focused exclusively on its objects, which are merely ideal, the a priori conditions of experience, on the other hand, concern our mode of cognition of real objects insofar as it presents their a priori components. It is nonetheless a form of “cognition,” with its own truth claim—not empirical, to be sure, but transcendental. And since, for Kant, truth means *adaequatio*, that is, the correspondence of cognition with its correlate,¹⁶ transcendental truth also involves such a correlate: possible experience.

In order to establish this correlation we can mention the two passages in the Analytic dealing with this specific truth claim. The first one appears, as we know, in the Schematism chapter: “transcendental truth, which precedes all empirical truth and makes it possible, consists in the general relation to this [the entirety of all *possible* experience].”¹⁷ The schemata are those products of the imagination that allow the pure concepts of the understanding to connect with possible experience. These a priori concepts thereby acquire their truth, that is, their objective validity, since from then on they have a correlate that can ground their claim to transcendental truth. The second passage is to be found in the Postulates of Empirical Thinking, to which we will return later on. It concerns the categories of relation, which gain their objective validity by their mere reference to experience in general:

¹⁴ See P 373: “the word ‘transcendental’...does not signify something passing beyond all experience but something that indeed precedes it a priori, but that is intended simply to make *knowledge of experience possible*” (my emphasis); Kant, *Prolegomena to Any Future Metaphysics*, trans. Lewis White Beck, in *Philosophic Classics*, ed. Forrest E. Baird and Walter Kaufmann, vol. 3, *Modern Philosophy* (Upper Saddle River, NJ: Prentice Hall, 1997), 581.

¹⁵ A96.

¹⁶ A58/B82.

¹⁷ A146/B185.

“one cognize[s] their objective reality [of the categories of relation], i.e., their transcendental truth, and, to be sure, independently of experience, but yet not independently of all relation to the form of an experience in general.”¹⁸

On the one hand, the transcendental principles laid out in the *Analytic* acquire their truth through their relation to possible experience; on the other hand, experience owes its own possibility to these a priori conditions. “Possible experience” thus has a twofold meaning: (1) experience is made possible by a priori conditions, and (2) possible experience confers validity on the entire transcendental apparatus.¹⁹

This correlation is difficult to grasp since it takes place at a virtual level; it has no anchorage point, so to speak. If the correlate of the a priori conditions of experience were a given object to which such cognition could correspond, the reader of the *Critique of Pure Reason* would have a much easier task. But as we have seen, this correlate is a mere possibility. There is no actual object nor any well secured knowledge to rely on. It goes without saying that Kant is fully aware of this situation, which is bound up with his way of proceeding in the *Critique*, as he will himself later admit in the *Prolegomena*. The *Critique* does not presuppose any “fact” whatsoever.²⁰ The only thing that is given is reason as a faculty of cognition, which contains (together with the pure forms of intuition) the a priori elements that make all cognition, and indeed all objects of cognition, possible. Beginning with those a priori elements, the first *Critique* adopts a synthetic and progressive procedure that leads to the possibility of

¹⁸ A221–2/B269.

¹⁹ A157/B196.

²⁰ See P 274 (Kant, *Prolegomena*, 526):

In the *Critique of Pure Reason* I have treated this question synthetically, by making inquiries into pure reason itself and endeavoring in this source to determine the elements as well as the laws of its pure use according to principles. The task is difficult and requires a resolute reader to penetrate by degrees into a system based on no data except reason itself, and which therefore seeks, without resting upon any fact, to unfold knowledge from its original germs.

On this topic see also Manfred Baum, “Die Möglichkeit der Erfahrung und die analytische Methode bei Reinhold,” in *Philosophie ohne Beynamen*, ed. Martin Bondeli and Alessandro Lazzari (Basel: Schwabe, 2004), 104–18.

experience. In contrast, using an analytic method, the *Prolegomena* of 1783 offers a more accessible presentation of transcendental philosophy beginning from two sciences: pure mathematics and pure natural science, both of which are already firmly established and aptly combined in Newtonian physics. A regressive analysis—from the given to its presuppositions—of these pure sciences thus allows a much easier access to their a priori conditions of possibility.

As a reminder, we may quote a passage from the Transcendental Methodology which highlights the difficulty of proving a transcendental principle when the regressive procedure of the *Prolegomena* is excluded, as is the case in the first Critique. What I have in mind is the end of the chapter on the Discipline of Pure Reason in Dogmatic Use, where Kant writes that the principle of causality can be proven apodictically, but only with regard to the experience that is thereby made possible. The proof therefore cannot take experience (nor any science) for granted and rely on it as a point of reference, since experience is first made possible by this very principle: the principle of causality “has the special property that it first makes possible its ground of proof, namely experience, and must always be presupposed in this [experience].”²¹

The A Priori Conditions of Possible Experience

Among the conditions of experience, Kant distinguishes the conditions of the possibility of experience from the conditions of its actuality. The first are a priori and manifestly receive the most attention in the *Critique* since they are specifically transcendental, while the others are a posteriori, that is, empirical. To be sure, together they constitute what makes experience “possible,” yet, in the narrow sense, it is the a priori conditions that are especially concerned with the “possibility” of experience.

The entire set of conditions for experience, namely the conditions of the possibility, of the actuality, and of the necessity of an

²¹ A737/B765.

appearance, are laid out systematically in the three Postulates of Empirical Thinking in General:

1. Whatever agrees with the formal conditions of experience (in accordance with intuition and concepts) is *possible*.
2. That which is connected with the material conditions of experience (of sensation) is *actual*.
3. That whose connection with the actual is determined in accordance with the general conditions of experience is (exists) *necessarily*.²²

As we can see, the first postulate concerns the formal aspect of experience, and the second concerns its matter, that is, “sensation,” which is for Kant “the sole characteristic of actuality,”²³ whereas the third, which deals with an “existence” that is at the same time recognized as necessary, is simply a combination of the first two postulates.²⁴ In this third case we can think of Kant’s famous example, “the sun warms the stone.” The *actuality* of the warmth of the stone (known through sensation) is *necessary* because it is the *effect* of the sun’s rays.

To be sure, the first postulate has to do with the conditions that make an “object” (*Ding, Gegenstand*)²⁵ possible, but since the Transcendental Deduction shows that experience and its objects share the same conditions of possibility,²⁶ the “formal” conditions in the first postulate are clearly the a priori conditions of possible experience in general. In other words, while the second postulate deals with the empirical conditions of experience, the first introduces its transcendental conditions, namely the pure forms of intuition and the pure concepts of the understanding. These a priori

²² A218/B265–6.

²³ A225/B273.

²⁴ According to Giuseppe Motta, the postulate of necessity has priority over the two others. See his “Qu’est-ce qu’un postulat? Considérations sur l’anti-constructivisme de Kant,” in *Kant et la science: La théorie critique et transcendantale de la connaissance*, ed. Sophie Grapotte, Mai Lequan, and Margit Ruffing (Paris: Vrin, 2011), 142–3.

²⁵ A220/B267.

²⁶ “The a priori conditions of a possible experience in general are at the same time conditions of the possibility of the objects of experience” (A111).

elements are the starting point of the Deduction, which endeavors to demonstrate the validity of these a priori formal conditions.²⁷ It goes without saying that these intuitive and conceptual elements are combined in the eight principles presented in the Transcendental Analytic. The principles show the interaction of these elements that, as a priori synthetic propositions, embody the formal conditions of the possibility of experience. Let us note that experience does not have to be *actual* at this stage, since these transcendental propositions acquire their objective validity simply by making experience *possible*.

Now in order to illustrate the finite character of Kant’s transcendental philosophy, we must consider more closely the nature of these transcendental conditions of experience. As we know, Kant divides his table of the principles of the understanding under two headings: mathematical and dynamical. Of these two classes, it is the dynamical principles that are the most relevant to my purpose here, but their special status can best be explained by first establishing a contrast with the mathematical principles, namely the Axioms of Intuition and the Anticipations of Perception. Obviously, these principles do not themselves belong to mathematics, yet they deserve this name insofar as they authorize the full application of the results of mathematics to experience. The reason for this is simple, but it should nevertheless be recalled: the synthesis of apprehension of the manifold in experience is, according to the Axioms of Intuition, the “same synthesis”²⁸ as the one at work in the quantitative synthesis of the homogeneous units of pure intuition in mathematics. This means that all the operations of quantification made a priori in arithmetic and in Euclidian geometry are automatically and universally applicable to experience, so that their objective validity is guaranteed from the start. It thus comes as no surprise when Kant argues that

²⁷ “There are two conditions under which alone the cognition of an object is possible; first, intuition, through which it is given, but only as appearance; second, concept, through which an object is thought that corresponds to this intuition...All appearances therefore necessarily agree with this *formal* condition of sensibility...The objective validity of the categories, as a priori concepts, rest on the fact that through them alone is experience possible (as far as the *form* of thinking is concerned)” (A92–3/B125–6, my emphasis).

²⁸ B203; see also A165–6.

these principles are “constitutive” of the objects of experience as far as their intuition is concerned. The quantification of extensive as well as of intensive magnitudes can be fully anticipated; for example, everything that geometry says about space as a formal intuition is by the same token valid for concrete experience. This is also what Kant has in mind when, speaking of intensive magnitude, he writes that the luminosity of the sun can be calculated a priori. To be sure, the sun and its light are not known a priori, but the calculus as such allows one to “construct” this magnitude “a priori.”²⁹

It is obvious that all the principles of the Transcendental Analytic are “necessary” conditions of the possibility of experience, but the mathematical principles are further declared “necessary” in their very exercise (*Ausübung*): their evidence is intuitive and they are a priori constitutive of the object of experience itself, at least as far as its anticipated magnitude is concerned.³⁰ On the other hand, the application of the dynamical principles, that is, the Analogies of Experience and the Postulates of Empirical Thinking in General, is said to be merely “contingent,” in the sense that, even though these principles anticipate experience, their exercise is “indirect” and “mediated.”³¹ The reason for this distinction is that the latter principles are concerned not so much with the mere intuition of the object as with its existence. And existence cannot be constructed a priori; it must first manifest itself empirically in order to be regulated by the dynamical principles.

Consider the dynamical principle of causality. According to this principle, once an appearance has presented itself through a sensation (for instance, a stone that becomes warm), this event can only be interpreted as an effect whose cause must be sought in a preceding moment in time, such as the emission of sunrays. This givenness of the event (the warmth of the stone), which escapes the control of the knowing subject, is

²⁹ “I would be able to compose and determine a priori, i.e., construct the degree of the sensation of sunlight out of about 200,000 illuminations from the moon. Thus we can call the former principles constitutive” (A178–9/B 221).

³⁰ A160/B199.

³¹ A160–1/B199–200. In their own specific way, all the principles of Kant’s table “anticipate” experience. See for example A246/B303, A762/B790.

what leads Kant to claim that the exercise of these dynamic principles is merely “contingent”: “the a priori conditions...of the *existence* of the objects of a possible empirical intuition are in themselves only *contingent*.”³² This comes somewhat as a surprise, because it seems to weaken the status of the dynamical principles, which are central for the possibility of experience. Not only is the principle of causality the one that recurs most often when Kant seeks to illustrate the role of the transcendental principles, but we also know, thanks to Paul Guyer,³³ that in the *Duisburg’sche Nachlass*, the Analogies of Experience, then called the “analogies of appearance,” were at the center of Kant’s early efforts to establish the conditions of the possibility of experience, whereas the mathematical principles seem to have been added only to the later versions of the table of the Analytic, in view of the publication of the *Critique*. Now the question is: What is so particular about the dynamical principles that warrants their being reduced to the status of mere contingency?

The Contingency of Some of the A Priori Conditions of Experience

We are facing a problem similar to the one we encountered concerning the concept of possibility. Like the latter, contingency belongs to the categories of modality, here as the counterpart of the concept of necessity. Yet the nominal definition of contingency does not reveal much, and so beyond the “logical” meaning of the word we have to search for some “real” meaning. The logical definition of contingency simply refers to something whose nonbeing does not imply contradiction.³⁴ For a more telling definition, for a “real” definition, we must turn to the way Kant applies it within experience. The details of this use are spelled out in the General Note on the System of the Principles added to the second

³² A160/B199, my emphasis.

³³ Paul Guyer, *Kant and the Claims of Knowledge* (Cambridge: Cambridge University Press, 1987), 27, 33, 35, 41–2.

³⁴ B290; A459/B487.

edition of the *Critique*.³⁵ There we learn that contingency as a category of modality can only be understood if it is linked with another set of categories, namely the categories of relation. Indeed, contingency is for human knowledge essentially a “relational” concept. It must be related to something else, because otherwise we would be left with an “absolute contingency,” which according to Kant is too “big” a concept for us.³⁶ This coincides with our earlier findings about the absolute possibility: it is the “most” that can be said on the possibility of a concept, and as such a finite mind cannot understand what it means. Let us recall that Kant refrains from deciding whether, for example, substance in experience is *in itself* contingent or necessary—we simply cannot know, since such a knowledge would call for a speculative use of reason.³⁷ Contingency then applies only to the accidents of appearances, since they can be interrelated. Kant gives the case of a body that was first in movement and has come to rest. As such, the mere change from movement to rest proves neither the contingency of this resting state nor of the movement observed earlier: these contrary states of the body occur at different times. To claim that the initial movement is contingent, we have to admit that at that initial time the body could just as well have been at rest. Yet this can only be made comprehensible to our finite knowledge if we presuppose that the reason why the body was in motion, rather than at rest, is that this motion results from some cause. And this is what Kant has in mind when he claims that contingency is comprehensible solely with the help of the categories of relation, in particular cause-and-effect: the movement of the body is a contingent state because, without the action of a cause, the body would have initially been at rest. Contingency is intrinsically linked with the concept of cause.³⁸

³⁵ B288–91.

³⁶ Kant, *Vorlesungen über Metaphysik (v. Schön)*, AK 28:499. See also Kant, *Preisschrift über die Fortschritte der Metaphysik*, AK 20:329–30.

³⁷ A635/B663; A227/B279.

³⁸ “That the proposition ‘Everything contingent must have a cause’ may be evident to everyone from mere concepts is not to be denied; but then the concept of the contingent is already taken in such a way that it contains, not the category of modality (as something, the non-existence of which can be thought), but that of *relation* (as something that can only exist as the consequence of

Can we now apply this reading, valid within the field of experience, to our prior question about the contingency of the transcendental principles that Kant calls dynamical? To be sure, the relation to be established here cannot be to a physical “cause,” as in the example above. But it should at least take the form of a “condition,” to use a more neutral term. If the Analogies of Experience are contingent in their application, this means that, as a priori conditions of experience, they are themselves “conditioned.”³⁹ And this is exactly what Kant claims in the passage where he declares these principles to be contingent: “the principles of the dynamical use, to be sure, also carry with them the character of an a priori necessity, but *only under the condition* of empirical thinking in an experience, thus only mediately and indirectly.”⁴⁰ While the mathematical principles could precede the material conditions of experience, pertaining as they do only to the form of an appearance or to the intensive magnitude of its matter, the dynamical principles on the other hand are in an essential manner related to the “empirical” conditions. This holds for the Analogies as well as for the Postulates of “Empirical” Thinking. They are conditioned by perception, because they are tied to the givenness of the object. This was already clear in the sentence quoted above on the contingency of these principles: they deal with the “existence of the objects of a *possible empirical* intuition.”⁴¹ Intuition here is not formal but clearly “empirical.” This empirical conditioning does not have to be actual, however; it merely has to be “possible.”

something else), and then it is, of course, an identical proposition: “What can only exist as a consequence has its cause” (B289–90, my emphasis). See also Giuseppe Motta, *Die Postulate des empirischen Denkens überhaupt* (Berlin: de Gruyter, 2012), 62.

³⁹ If the General Note on the System of Principles establishes that we cannot understand contingency except when the contingent event is to be explained through a “cause” that precedes it in time, then Kant cannot have recourse to the schematized category of causality in his philosophical investigation, which describes the conditions of experience from without. Nevertheless, the restriction imposed upon contingency is maintained insofar as reference is still made to a “condition” (even if it is not a cause in time).

⁴⁰ A160/B199–200, my emphasis.

⁴¹ A160/B199, my emphasis.

The following conclusion can be drawn from the contingent intervention of the dynamical principles: among the a priori *conditions* of the possibility of experience there are some, indeed the most important ones, that are themselves *conditioned*. And this result does indeed conform with the “real” meaning of contingency according to Kant: “all contingency is only possible in a conditioned manner.”⁴² This plainly amounts to saying that the transcendental point of view in the *Critique* is not absolute, but relative. Being contingent in their application, the dynamical principles are dependent upon something else: the material (empirical) conditions.

Naturally both the formal and the material conditions are combined for the production of experience, but it is important to specify the terms of this joint contribution. In his lectures on metaphysics, Kant deals repeatedly with the question of multiple causes contributing to a single effect (in our case: experience). He calls them *concaussae* (or *Mitwirkungen*) and suggests that there are two ways for them to bring about their effect: either they run parallel and remain merely coordinated, or they are subordinated.⁴³ If we apply this model to the transcendental production of experience, we are compelled to conclude, after what we have seen, that the dynamical principles are subordinated to the possibility of the material conditions, at least if they are to be a priori cognitions at all. These a priori synthetic propositions are conditioned cognitions, which can in no way lay claim to the status of the unconditioned and which therefore do not have to be implemented by a transcendental subject conceived as an intuitive understanding.

What is said—and what can be said—of the transcendental subject is restricted to its role in the constitution of experience. The spontaneity of the understanding, which culminates in the unity of apperception, is known only through the way it “affects” inner sense. In this case inner sense contains the reflection of this activity, which consists essentially in combining

⁴² Kant, *Vorlesungen über Metaphysik (Dohna)*, 1792–93, AK 28:647.

⁴³ Kant, *Vorlesungen über Metaphysik (Mrongovius)*, 1782–83, AK 29:844 : “Viele Ursachen, sofern sie zu einem caussato gehören, heissen concaussae, die sind entweder sibi subordinatae, wenn eine vermittelt der andern caussa caussati ist—oder coordinatae, wenn keine als caussa remota, sondern alle als immediate anzusehen sind.” Karl Ameriks has drawn attention to these multiple causes in *Interpreting Kant’s Critiques* (Oxford: The Clarendon Press, 2003), 155 n.42.

representations.⁴⁴ The empirical representations are at first present in their merely subjective order, and the role of the understanding through its spontaneity is to introduce an objective order among them. For example, if the perception of the warm stone comes first and then afterward the perception of the rays of the sun, it would be erroneous to claim, in accordance with the subjective temporal succession, that the first perception is the cause of the second and to say that the warm stone has produced the hot sunrays. The understanding simply restores the objective order of things.

But the thinking subject thereby runs up against the limits of its spontaneity. Such an activity is limited to the combination of the representations in inner sense. The spontaneity of the understanding does not however *produce* the manifold of these representations, since it would then be an intuitive understanding. If Kant’s transcendental investigation aims at establishing the limits of human knowledge, these limits must also be recognized within the transcendental apparatus: “that understanding through whose self-consciousness the manifold of intuition would at the same time be *given*, an understanding through whose representation the objects . . . would at the same time *exist*, would not require a special act of the synthesis of the manifold for the unity of consciousness, which the human understanding, which merely thinks, but does not intuit, does require.”⁴⁵ In other words, an understanding that would directly intuit the existence of its object would not be conditioned. It would be totally independent and self-sufficient. It would produce its object of knowledge without any further condition, that is, without external conditions.

We have seen that the material conditions are precisely what show the limits of the a priori conditions of experience. If, on the one hand, inner sense can be affected “from within” by the spontaneity of the understanding, it can also be affected “from without,” via outer sense, by perception, which represents a radical type of otherness vis-à-vis the transcendental subject.⁴⁶

⁴⁴ “I do not see how one can find so many difficulties in the fact that inner sense is *affected* by ourselves. . . . In such acts the understanding always determines the inner sense, in accordance with the combination that it thinks, to the inner intuition that corresponds to the manifold in the synthesis of the understanding” (B156–7 n, my emphasis).

⁴⁵ B138–9, my emphasis. See also B135.

⁴⁶ B156.

Kant sometimes calls it the “transcendental object,” but it is best known as the notorious “thing in itself.” This is the sphere where the passivity of the subject begins, where receptivity takes on its full significance.

I know all too well that the thing in itself is a hotly disputed topic, but I simply want to stress that Kant’s positing of the thing in itself does not represent a leap into the unconditioned. On the contrary, the thing in itself is the ultimate recognition of the conditioned character of the knowing subject and of his or her knowledge. This second kind of affection of inner sense (via outer sense) cannot be attributed merely to the appearances, nor can it be regarded as a mere epistemic requirement, that is, an object of thought. It is an ontological statement that is constitutive of Kant’s investigation of the limits of human knowledge. As Karl Ameriks has argued, there has to be a legitimate place for “transcendental affection” in Kant.⁴⁷ For instance, Kant is aware that the “I think” of the transcendental deduction contains an existential statement, although he refrains from exploiting it in a Cartesian fashion and cautions that nothing can be known about the nature of the transcendental subject.⁴⁸ The Dialectic will remind us that it is even impossible to say that it is a substance.⁴⁹ It is nevertheless admitted as existing. Now it should be possible to say the same of the counterpart of the transcendental subject: the thing in itself. Nothing can be known of its nature, but its specific kind of affection of sensibility is nonetheless essential for the general possibility of experience.⁵⁰

⁴⁷ See Ameriks, *Interpreting Kant’s Critiques*, 157. See also Nicholas Stang, “Did Kant Conflate the Necessary and the A Priori?,” *Nous* 45, no. 3 (2011): 467 n.16. In this note Stang explains that there are so many passages in Kant referring to the affection by the thing in itself that it would be disingenuous to deny them.

⁴⁸ See B157, B277.

⁴⁹ B422. On Kant’s recourse to the term “*das Substantiale*” in order to explain the dialectic production of the “fiction” of a transcendent spiritual substance, see my “Die Entstehung der Illusion in den Paralogismen,” in *Über den Nutzen von Illusionen: Die regulativen Ideen in Kant’s theoretischer Philosophie*, ed. Bernd Dörflinger and Günter Kruck (Hildesheim: Olms, 2011), 47–58.

⁵⁰ Here arises a similar problem to the one we encountered in the case of contingency. Kant frequently uses the word “ground” (*Grund*) in order to explain the affection stemming from the thing in itself. But sometimes he uses the word “cause” (*Ursache*), as Aenesidemus-Schulze noted. It goes without saying that this dynamical category is not schematized here and thus does not

Conclusion

We have raised the question as to whether it was legitimate for Kant’s transcendental discourse to use certain modal categories beyond the realm of empirical knowledge (experience), that is, outside of the only type of cognition in which a priori synthetic propositions have access to intuition, formal as well as material. To this question we might answer that the modal categories in Kant’s transcendental discourse are ultimately focused on experience itself as a whole. In my view, this reference to possible empirical cognition is precisely what justifies the use of these modal categories and what confers upon them a claim to truth in Kant’s transcendental discourse, although, to be sure, this claim to objective reality is indirect. We are far from the typically “transcendental” use of the categories in the Dialectic. After all, experience is not an absolute possibility, much less an unconditioned necessity, like the dialectical idea of God, for example. We are rather led to conclude with Kant that possible experience is in the end something “entirely contingent,”⁵¹ since it depends on a broad set of conditions. Furthermore we have discovered that the most important among the transcendental conditions are themselves contingent in turn. Unlike freedom, say, which is construed in the thesis of the third Antinomy as an absolute causality, they are not unconditioned conditions. Finally, the focal point of the transcendental conceptual conditions, the “I think,” is not declared in the Analytic to be a noumenal substance, which is what the dialectical

apply to an appearance within experience. It is instead used to articulate the conditions of possible experience as a whole. But, as opposed to the dialectical use of the category, it still respects the constraints stated in this principle: at first only the effect is known and the principle simply stipulates that there must be *some cause or other*, which is at that stage totally “indeterminate” (A179/B 222; A199/B244). The only thing that is certain, according to the principle of causality, is the cause’s existence, since the principle cannot anticipate anything of its essence. Yet the unschematized category of causality used by Kant for the thing in itself clearly complies with the indeterminateness mentioned in the principle. And again, its ultimate justification is the possibility of experience, which by definition relates such a cause to the possibility of an “empirical intuition.” See my “Kant and the Problem of Affection,” *Symposium: The Canadian Journal of Continental Philosophy* 8, no. 2 (2004): 275–97.

⁵¹ A737/B765.

reasoning in the Paralogism of Pure Reason tries to establish. On the contrary, the “I think” is mobilized solely for its contribution to making experience possible.

Claude Piché is Professor of Philosophy at the University of Montreal. Author of *Kant et ses épigones: Le jugement critique en appel* (Librairie philosophique J. Vrin, 1995) and *Das Ideal: Ein Problem der Kantischen Ideenlehre* (Bouvier, 1984) as well as numerous articles, he specializes in Kant, Kantianism, and German idealism.

2

Plato and Kantian Transcendental Constructivism

Tom Rockmore

The meaning of “German idealism” is, like the meaning of “idealism,” unclear. There is nothing resembling agreement among observers about either term. In a recent book, I argue that Kant and the post-Kantian German idealists share a common interest in solving or resolving the cognitive problem in all of its many forms along constructivist or Copernican lines.¹

Philosophy and the History of Philosophy

If for no other reason, the relation of philosophy to its past is significant for understanding and evaluating it. It must be very rare that wholly new themes, without precedent of any kind, arise in the debate.

¹ See Tom Rockmore, *German Idealism as Constructivism* (Chicago: University of Chicago Press, 2016).

T. Rockmore (✉)

Department of Philosophy, Peking University, Beijing, China

It seems obvious that most philosophical theories are formulated to respond to problems, difficulties, and concerns transmitted by the prior debate. This suggests philosophy builds on the prior debate, or on the history of philosophy. Yet, since this question has never been settled, every generation apparently needs to fight this battle anew.

The history of philosophy has been out of fashion for centuries. In modern times, though there are exceptions, an approach to philosophy, particularly the critical philosophy, through the history of the tradition, is unusual. Modern philosophy often emphasizes the independence of philosophy with respect to its history. Thus Descartes thinks of earlier theories as comprising a series of mistakes. This point is made in different ways by numerous later thinkers. They include the early Wittgenstein, who thinks of philosophy as depending on the philosophical misuse of language; Quine, who draws attention to the distinction between those interested in philosophy and those interested in the history of philosophy; Husserl, who is concerned finally to make a true beginning; Heidegger, who seeks to return back behind the tradition to recover the central question of philosophy, and so on. These and others think the case still needs to be made for considering philosophy against the historical background.²

Kant's Critical Philosophy and the History of Philosophy

The relation of the critical philosophy to the history of the discipline is unclear. Kant is apparently of two minds about this relation. He suggests his position is both independent of, as well as dependent on, the tradition in various ways. The critical philosophy

² See, for discussion, Konrad Cramer, "Das philosophische Interesse an der Geschichte der Philosophie," in *Subjektivität und Autonomie*, ed. Stefan Land and Lars-Thade Ulrichs (Berlin: de Gruyter, 2014), 33–50.

is independent of the history of philosophy in virtue of its transcendental status, its a priori nature, its supposed unrevisability, and so on. As a transcendental theory the critical philosophy lays claim to provide the only possible approach to cognition. As an a priori theory, it formulates a position independent of time and place. As a supposedly unrevisable position, Kant's theories, like Ozymandias's statue, are intended to stand forever.

The same thinker who suggests his position is independent of the history of philosophy, hence ahistorical, appears to contradict this view in various ways. They include his reference to generalizing Hume's problem, his reliance on Leibniz's logical view of causality, his praise for Wolff, his restatement of a Cartesian conception of the subject, his concern with whether various types of cognition are on the secure road of science, and so on. These and similar indications point to Kant's effort to formulate the critical philosophy not in ignoring but rather in building on the prior tradition.

Kant and Ancient Greek Philosophy

It is sometimes noted that Kant reacts to the Greek tradition in formulating the critical philosophy. The Verona school, for instance, emphasizes the influence of Aristotle on Kant.³ Another approach is suggested in Kant's important remark that it is not rare that we know an author such as Plato better than he knows himself.⁴ This suggests at least four consequences: first, Plato has a philosophical view; second, Kant possibly knows Plato's view better than Plato; third, it is by inference possible to know the critical philosophy better than its author; and, fourth, despite the famous reference to Hume, it is at least plausible that Kant formulates the cognitive problem as well as his

³ See, for example, Marco Sgarbi, *Kant on Spontaneity* (London: Bloomsbury, 2012).

⁴ Immanuel Kant, *Critique of Pure Reason*, trans. and ed. Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press, 1998), A314/B370.

response to it in Platonic terms, to which the critical philosophy can be understood as a response.

I think we should take this Kantian hint seriously. A different version of this suggestion was made more than a century ago by Natorp, following Cohen,⁵ in his Kantian interpretation of Plato's theory of forms (or ideas). In *Platos Ideenlehre* (1903), Natorp develops a "critical" interpretation of the notorious theory of forms as well as an argument for the order of the dialogues in the context of an "Introduction to Idealism." Natorp, who thinks Plato was misinterpreted since Aristotle, denies the familiar interpretation of Platonic ideas or forms as things or substances. According to Natorp, Platonic forms are to be understood as laws or methods, and thus as foundational for science in depicting Plato as the founder of critical idealism.

My view is related to but different from Natorp's belief that later Kantian idealism builds on Plato's theory of forms. The term "representation" is routinely understood in different ways. I will be suggesting that, if we comprehend "representation" as a form of cognition, then Kant can be read as accepting, not the Platonic theory of forms, but rather the associated Platonic interdiction of cognitive representationalism, or the view that we can correctly represent mind-independent reality, but rather the contrary view that we do not and cannot cognize reality in formulating a nonrepresentational, constructivist approach to cognition. The difference is roughly, as Kant points out in the Copernican revolution, between making cognition depend on the grasp of a mind-independent object, whose possibility Kant denies, and making the object rather depend on the subject, which Kant suggests as an experiment.

I should make it clear that I do not intend to depict Kant as a late Platonist or conversely Plato as an early Kantian. Kant's relationship to

⁵ Cohen was very interested in Plato, whom he understood as an early idealist thinker. See, for example, Hermann Cohen, "Die Platonische Ideenlehre psychologisch entwickelt," *Zeitschrift für Völkerpsychologie und Sprachwissenschaft* 4 (1866): 403–64; and "Platons Ideenlehre und die Mathematik," in *Rectoratsprogramm der Universität Marburg* (Marburg: Elwertsche, 1878).

Plato is unclear and controversial for various reasons. It is, for instance, unclear whether Kant ever read Plato or rather mentions him on the basis of indirect knowledge only. Yet this is not surprising since we also do not know how well Kant knew English,⁶ nor which of Hume's writings he in fact read.

Parmenides and the Cognitive Problem

We can, for present purposes, draw attention to a triple distinction between Parmenides' setting of the cognitive problem, Plato's relation to Parmenides, and Kant's relation to Plato. The general problem concerning Parmenides, Plato, and then Kant can usefully be raised in terms of realism, which has long driven the Western cognitive debate. Realism is an ontological theme central to Western epistemology. All conceptions of knowledge are realist and hence lay claim to grasping the real, however it is understood. Realism includes artistic, social, scientific, metaphysical, and other varieties.

Metaphysical realists believe there is a way the world is and that we can accept as our standard nothing less than a cognitive grasp of the real. This view appears perhaps for the first time in Parmenides' poem. At B 8.34, in writing "*to gar auto noein estin kai einai*,"⁷ he points toward what later becomes metaphysical realism by opting for identity as the standard of knowledge. This interpretation is supported by textual analysis. Thus Burnyeat, who thinks idealism is a specifically modern doctrine, believes Parmenides holds that thought refers to being.⁸

Parmenides apparently asserts that cognition depends on identity between subject and object, knower and known. Various types of identity can be distinguished. Frege stresses semantic identity in claiming that the morning

⁶ See, for discussion, Sanford Budick, *Kant and Milton* (Cambridge: Harvard University Press, 2010).

⁷ Diels, H. & Kranz W., eds., *Die Fragmente der Vorsokratiker* (Berlin, 1952), 28 B 3, Clem. Alex. Strom. 440, 12; Plotinus, *Enneads* 5, 1, 8.

⁸ "But the fragment (frag. 3) which was once believed, by Berkeley among others (Siris §309), to say that to think and to be are one and the same is rather to be construed as saying, on the contrary, that it is one and the same thing which is there for us to think of and is there to be: thought requires an object, distinct from itself, and that object, Parmenides argues, must actually exist." M. F. Burnyeat, *Explorations in Ancient and Modern Philosophy* (Cambridge: Cambridge University Press, 2012), 255.

star and the evening star have different meanings but the same reference. Numerical identity is the sense in which a given thing is self-identical. For instance, the feather pen Krug employed to criticize Hegel is in this sense identical to his writing instrument. Qualitative identity, which refers to the way in which two or more things share a property, is illustrated in the Platonic theory of forms (or ideas). Identity in difference, which is neither numerical nor qualitative, is a metaphysical relation brought about by the subject in creating a unity between itself and the object it “constructs.”

In different ways what I am calling the Parmenidean view echoes through the tradition in the form of an identity in difference of thought and being. Much later in the German idealist tradition this Parmenidean identity becomes the identity of identity and difference. Thought and being are obviously not the same, since being, or what is, is independent of thought about it. But from the Parmenidean perspective “to know” means that “thought grasps mind-independent being.” Since, according to Parmenides, cognition depends on an identity of thought and being, we can infer that a necessary condition of cognition is an identity of thought and being. This point can perhaps be stated more precisely as the identity of thought that grasps, hence cognizes, mind-independent being as well as being that differs from thought, or difference.

The identity of identity and difference, which is identified with German idealism, only becomes explicit at the time of Hegel. Yet it is at least implicit throughout the Western philosophical debate on knowledge since the early Greek tradition. This identity is featured, for instance, in metaphysical realism, which echoes through the entire Western tradition up to the present day. The claim to know is routinely understood as a claim to grasp not what one thinks is the case but rather what in fact really is. Since Western philosophy originated in ancient Greece, it has steadily examined different cognitive strategies for what is now called metaphysical realism. The history of the philosophical debate on knowledge consists in a long, varied, often ingenious series of efforts to demonstrate the claim to know the mind-independent world. Yet other views of knowledge, including those that restrict cognitive claims merely to phenomena, and which are featured throughout German idealism, simply give up any form

of the ancient effort to know reality while maintaining the claim for the identity of identity and difference.

Metaphysical Realism and Constructivism

We can summarize these remarks about Parmenides in two points. His suggestion that the cognitive problem requires an identity of subject and object, thought and being, or again knower and known, suggests two possible solutions: either one must grasp mind-independent reality as it is, or one must construct what, since it is given in experience, is not mind-independent reality, not the real as such, but, since it is dependent on experience, only the real for us.

These two solutions point in opposing directions. The first solution points toward the metaphysical realist view that a necessary condition of knowledge is to grasp the mind-independent external world not merely as it appears but as it in fact is. The claim that to know is to know reality, or the real, is arguably the main strand in the debate on cognition since its origin in ancient Greece. The counter-claim that we do not and cannot know the mind-independent world as it is, in other words that metaphysical realism is an impossible requirement for cognition, points in the other direction, towards cognitive constructivism. Metaphysical realism has over the centuries been a main element in cognitive theory of the most varied kinds, which, since early Greek thought, continues to rely on the view that to know requires us to know the world as it is. This view, which, like ice cream, comes in many flavors, remains as popular now as in the ancient world. It seems current interest in scientific realism, social realism, and so on are merely variations on the metaphysical realist theme.

Parmenides is perhaps the initial “modern” figure in the tradition. The metaphysical realism suggested by his position continues to influence the cognitive debate. Unlike metaphysical realism, constructivism is a second-best view, a view toward which one turns if it seems that metaphysical realism fails to avoid cognitive skepticism. Cognitive constructivism is any form of the view that we do not and cannot cognize the mind-independent world as it is, since we can only know that we cognize what we in some

sense construct, produce, or make. This view arises in ancient mathematics, in the Euclidean view that the construction of a geometrical figure with a straight edge and compass is proof of the existence of the entire class.

Parmenides and Plato

The Parmenidean view that cognition requires the identity of thought and being points in two different directions: toward metaphysical realism or, on the contrary, toward cognitive constructivism.

Plato is the single most influential proponent of the Parmenidean view that to know is to know the mind-independent world as it is. The theory of forms suggests there is direct, intuitive knowledge of the real. According to Plato, some selected individuals, on grounds of nature and nurture, are able to intuit or even literally to see the mind-independent reality. We do not know and cannot determine if Plato accepts any form of the notorious theory of forms, which is routinely attributed to him, but we do know that he rejects a causal analysis of knowledge. Thus he accepts the ontological view that forms cause appearances but rejects the backward cognitive inference from appearances to reality. An example might be an inference from a table, which for Platonism is the effect of which the form of the table is the cause, to the form of the table. Plato rules out this kind of inference, though he suggests that philosophers can directly intuit reality. Since he thinks that only philosophers can know, he famously excludes artists and poets, who do not and cannot know, from the city-state.

It is unclear what Plato's view is or even if he has a position in a modern sense. He could be saying that there is knowledge since philosophers in fact do know reality. Or he could be saying that if there is knowledge then it must be the case that philosophers in fact know reality.

Plato's influential support of the Parmenidean suggestion that cognition requires knowledge of reality continues to echo through the tradition. Examples include the Cartesian view that there are clear and distinct ideas about the world, the Lockean view that simple ideas match up one-to-one with the world, and, more recently, Davidson's claim that in giving up the dualism between scheme and world we come

into direct touch with the latter, or Brandom's nearly identical suggestion that reality makes our views of electrons or aromatic compounds true or false. Each of these thinkers lays claim to know the mind-independent external world as it is.

Kant and Plato

It has already been noted that Plato makes two crucial cognitive claims: we know or at least some of us know reality through direct intuition, and we do not and cannot know reality through a causal analysis since we cannot justify a backward causal inference from appearance to reality.

Modern philosophy, with exceptions, mainly turns away from cognitive intuition and toward cognitive representation in rehabilitating against Plato the reverse cognitive inference he rejects. Kant agrees as well as disagrees with Plato at two crucial points as concerns representation and intuition. Plato, we recall, denies representation in favor of intuition. Kant defends what initially seems like an ambiguous position. Three points are important. First, he disagrees with Plato in denying intellectual intuition. Second, he persistently features representationalist terminology even after he may have turned away from a representationalist approach to cognition. Third, he agrees with Plato in denying the backward anti-Platonic inference, hence in denying representation of the real, or in his terminology the thing in itself or noumenon. The ambiguity lies in the apparent conflict between the representationalist terminology and the denial of cognitive representationalism.

Criticism of a Copernican Reading of the Critical Philosophy

Since Kant denies intellectual intuition, he requires a different justification for cognitive claims. His positive argument for knowledge lies in the claim that we cognize only what we in some sense construct. This is the central insight of the famous Copernican revolution, which, if this approach is correct, lies at the heart of the critical philosophy. The Copernican revolution, a term Kant never uses to designate his position,

is a modern form of constructivism. This reading of the critical philosophy is controversial for a number of reasons. They include representation, Kant's knowledge of Copernicus, his link to Plato, and the extent to which we understand Kant.

We can begin with Kant's relation to representationalism, which is arguably the favored modern cognitive strategy. Early and late representationalist terminology pervades Kant's texts. Further he seems to feature representationalism in the famous Herz letter early in the critical period, where he asks: "What is the ground of the relation of that in us which we call 'representation' to the object [*Gegenstand*]?"⁹ This implies that he, like many other modern thinkers, is an epistemic representationalist. Yet he also explicitly states that representation cannot be defined at all.¹⁰

A constructivist approach to Kant links the critical philosophy to Copernican astronomy. But the most thorough study we possess of this question indicates Kant may never have read Copernicus at all.¹¹ Yet that is perhaps not important, since we also do not know to what extent he was familiar with Hume¹² or Plato.¹³

A constructivist reading of the critical philosophy casts light on our understanding of the critical philosophy. Kant is closely studied in an enormous and growing debate. Yet it is possible, since the Copernican revolution in his thought is little studied,¹⁴ and there

⁹ Immanuel Kant, *Philosophical Correspondence, 1759–99*, trans. and ed. Arnulf Zweig (Chicago: University of Chicago Press, 1967), 71.

¹⁰ In the *Dohna-Wundlacken Logic*, presumably based on lectures given in the 1790s, hence in the critical period, he states representation "cannot be explained at all." Immanuel Kant, *Lectures on Logic*, ed. J. Michael Young (New York: Cambridge University Press, 1992), 440.

¹¹ See Hans Blumenberg, "What Is 'Copernican' in Kant's Turning?," in *The Genesis of the Copernican World*, trans. Robert M. Wallace (Cambridge, MA: MIT Press, 1987), 595–614.

¹² See, for a thorough study of the relation of Kant to Hume, Paul Guyer, *Knowledge, Reason, and Taste: Kant's Response to Hume* (Princeton: Princeton University Press, 2013).

¹³ According to Kühn, Kant "seems" to have read Plato. See Manfred Kühn, *Kant: A Biography* (New York: Cambridge University Press, 2011), 370 n.19.

¹⁴ It is, for instance, not mentioned at all in a recent, detailed study of Kant's metaphysical approach to Newtonian mechanics and modern science in general. See Michael Friedman, *Kant's Construction of Nature: A Reading of the "Metaphysical Foundations of Modern Science"* (New York: Cambridge University Press, 2013).

is no agreement about its significance, that we do not understand the critical philosophy.

This is hardly implausible. It is obvious that the process of coming to grips with a great thinker is extremely lengthy, in some cases extending over hundreds of years. Anyone understood in his own time presumably has nothing of deep interest to communicate. It is possible that after several hundred years we still have not understood Kant.

Kant and Cognitive Constructivism

I have not argued that Kant is a Platonist. I have also not argued that Kant builds on Plato. I have rather argued that he builds on his rejection of Plato, more precisely on his refusal of the Platonic view of direct, intuitive knowledge of reality as a crucial factor in his Copernican revolution, or constructivist approach to cognition. Kant does not invent this approach, which comes into modern philosophy from ancient mathematics independently through Hobbes and Vico, possibly others. Kant rather reinvents cognitive constructivism as the basis of his effort to solve the cognitive problem.

A form of constructivism is advanced by Kant in his view of the cognitive object as constructed, hence cognizable, by the subject through its interaction with contents of the sensory manifold and the structures of the understanding. The transcendental deduction, which is often understood as a *quid juris*, is in part that, but above all an account of the general conditions of the construction, not of reality, but rather of the cognitive object.

It helps to understand German idealism as united through an effort by different hands to arrive at a plausible form of the constructivist view, which is initially formulated by Kant and then later pursued by Fichte and Hegel. It further helps to understand that, with respect to this criterion, Schelling falls outside the German idealist concern with constructivism.

The constructivist thesis still further helps us to understand the vexed relation between idealism and realism, which is central to the Marxist debate.

According to G. E. Moore, who relies on a misreading of Berkeley, idealism is a denial of the existence of the external world. Constructivism advances a claim, not for the construction of reality, which would be absurd, but for the construction of the contents of consciousness, for what we experience, not for what supposedly might exist independently of us.

On the Double Aspect Thesis

A constructivist interpretation of the critical philosophy is contradicted by the so-called double aspect thesis. This thesis, which is very popular at present, is the basis of Henry Allison's influential defense of Kantian transcendental idealism. The double aspect thesis interprets the critical philosophy through a twofold metaphysical commitment, which can be summarized thus: first, the mind-independent world affects the subject; and second, the affect, or result, and the cause, that is, the thing in itself, or again the noumenon, are two aspects of the same thing.

The double aspect thesis, for which there is ample textual evidence in Kant's writings, is anticipated in the modern tradition by Spinoza, Schelling, and more recently Heidegger. Spinoza famously insists, without justification of any kind, that thought and being run parallel. To Fichte's dismay, Schelling seeks to justify transcendental philosophy through its supposed parallel in philosophy of nature. Heidegger, who suggests his phenomenological ontology extends the critical philosophy, contends that being shows itself, hence that we know the world as it is.

The double aspect thesis is an extreme form of the modern causal theory of perception. It is extreme in that through the proposed rehabilitation of the anti-Platonic inference from effect to cause, it supposes that the effect correctly represents, or again is the faithful appearance of, the cause. In other words, according to this thesis, cause and effect are in principle identical.

This view is problematic. It merely asserts but does not demonstrate that noumena manifest themselves as appearances, in short that we know reality. Yet though this assertion is often made, it has never been

justified. It is because Kant thinks this inference is unjustified that he adopts the constructivist view, which is incompatible with representation of any kind, that we cognize only what is constructed by the subject. Yet if what we perceive and know is constructed by the subject, then we cannot infer from representations or appearances to noumena or things in themselves. Though Kant arguably never doubted the existence of the mind-independent external world, he clearly also holds that we can make no cognitive claims, none at all, about reality.

Conclusion: Plato and Kantian Constructivism

In this chapter I have examined the critical philosophy in relation to Plato. I have argued five points. To begin with, Parmenides influentially formulates the canonical criterion of knowledge at the dawn of the Western tradition. Second, Plato denies the backward inference from effect to cause, hence denies a representational approach to cognition. Kant further follows Plato in denying representation. Kant, unlike Plato, also denies cognitive intuition in opting for the Copernican turn or constructivism. I have finally argued that the dual aspect thesis is incompatible with Kantian constructivism.

I come now to my conclusion. The Platonic theory of forms suggests the impossibility of a backward inference from effect to cause while basing cognitive claims on intellectual intuition. The modern debate on knowledge often turns on denying intellectual intuition while in its place rehabilitating the backward anti-Platonic cognitive inference, in claiming, or at the very least in taking as the cognitive criterion, a cognitive grasp of reality.

In this respect, Kant is both modern and not modern at all. Like many other modern thinkers, he denies intellectual intuition. But unlike many others of the modern period, he denies as well the backward anti-Platonic inference from effect to cause. His Copernican revolution is an anti-Platonic effort to justify cognition, not of mind-independent reality, which does not appear and cannot be known, but rather of the empirically real that we construct as a necessary cognitive condition. Now, as before Kant, metaphysical realism remains a favored theme of the modern debate.

Yet as Kant points out, there has never been any progress toward this goal; and there has been none since Kant. Hence, short of epistemic skepticism, the only plausible approach lies, as he clearly saw, in the assumption that we construct what we know. In this respect, Kant was and perhaps still is ahead of his time, since he was, dare we say, postmodern.

Tom Rockmore is Distinguished Professor Emeritus at Duquesne University, as well as Distinguished Humanities Chair Professor at Peking University. He has strong interests throughout the history of philosophy and defends a constructivist view of epistemology. The philosophers whom he has studied extensively are Kant, Fichte, Hegel, Marx, Lukács, and Heidegger. Among his published books are *Kant and Phenomenology* (University of Chicago Press, 2011), *Kant and Idealism* (Yale University Press, 2007), *In Kant's Wake* (Wiley & Sons, 2008), and *Hegel, Idealism, and Analytic Philosophy* (Yale University Press, 2005).

3

Kant and Fichte on the Notion of (Transcendental) Freedom

Violetta L. Waibel

In the history of philosophy Johann Gottlieb Fichte can be considered the one philosopher other than Sartre who thought most emphatically about the notion of freedom. In a draft of a letter to the poet Jens Immanuel Baggesen written in April 1795, Fichte called the *Wissenschaftslehre* (Science of Knowledge) “the first system of freedom,” since it liberated man from the “bonds of the things in themselves and from external influences that have more or less exerted control over him in all systems created so far, even in Kant’s, and defines him in the first principle as an independent being.”¹ As Fichte wrote in his draft of a letter to Friedrich

¹ GA 3/2, no. 282, trans. Susanne Costa-Crivdic. See also Violetta L. Waibel, “Das ‘System der Freiheit’ und die ‘Feßeln der Dinge’: Fichtes Begründung der Gegenstandskonstitution (1794/95),” in *System der Vernunft: Kant und der Deutsche Idealismus*, vol. 2, *Kant und der Frühidealismus*, ed. Jürgen Stolzenberg (Hamburg: Meiner, 2007), 103–28. For another interpretation of Fichte’s “first system of freedom,” see Günter Zöller, “Das ‘erste System der Freiheit’: Fichtes neue Darstellung der

V.L. Waibel (✉)

Department of Philosophy, University of Vienna, Vienna, Austria

August Weißhuhn in August or September 1790, it was Kant himself who awakened him from his determinist, dogmatic slumber: “I have been living in a new world ever since I read the *Critique of Practical Reason*. Propositions which I thought could never be proved wrong, have been proved wrong to me. Things that I believed could never be proved to me like the notion of absolute freedom, of duty etc. have been proved to me, and I feel all the happier for it.” He writes further in the same letter, “I have thrown myself completely into Kantian philosophy, at the beginning because of necessity; I had to give a lesson about the *Critique of Pure Reason*. But after my encounter with the *Critique of Practical Reason* because of sheer delight.”²

It is therefore all the more remarkable that Fichte takes the notion of freedom much further than Kant. Kant does, however, distinguish between negative and positive freedom in his *Groundwork for the Metaphysics of Morals*, published in 1785, and in his *Critique of Practical Reason*, published in 1788. Negative freedom implies that humanity is independent of nature, objects, and inclinations.³ On further examination, however, we can see that for Kant freedom only has reality if a rational being is also a moral being. Only the one who acts morally and autonomously in the full meaning of that concept is truly free. In Kant’s critical philosophy freedom and autonomy are inseparably combined. He writes in the *Groundwork*:

As a rational being, and thus as a being belonging to the intelligible world, the human being can never think of the causality of his own will otherwise than under the idea of freedom; for, independence from the determining causes of the world of sense . . . is freedom. With the idea of freedom the concept of *autonomy* is now inseparably combined, and with the concept of autonomy the universal principle of morality, which in

Wissenschaftslehre (1795–1798),” in *System und Kritik um 1800*, ed. Christian Danz and Jürgen Stolzenberg (Hamburg: Meiner, 2011), 13–28.

² GA 3/1, no. 63, trans. Susanne Costa-Crivdic. Cf. Wilhelm G. Jacobs, *Johann Gottlieb Fichte: Eine Einführung* (Berlin: Suhrkamp, 2014), 12–14.

³ GMM 94–5, AK 4:446–7; cf. CPR 177–9, AK 5:47–8.

idea is the ground of all actions of *rational beings*, just as the law of nature is the ground of all appearances.⁴

Fichte, however, develops a theory of freedom in his *System of Ethics*, published in 1798, that introduces three stages of an ever more complex notion of freedom. These three stages are (1) thinking in concepts, which is what sets humanity apart from natural drives; (2) intentional, purposeful thinking, decision-making, and acting; and (3) acting morally in the full sense of autonomous reason.

The names of these stages alone make us wonder why Kant had not included them in his system as well. Indeed, it must be said that Fichte had not invented these different stages in the manifestations of freedom as such. The facts underlying these manifestations also have their place in Kant's Critiques. But whereas Fichte recognizes different forms of human freedom on all three levels, Kant only explicitly defines the third stage as an expression of freedom. Does this mean that Kant does not consider the expression of spontaneity, on the one hand, and planning, decision-making, and purposive rational acting, on the other, as manifestations of human freedom?

After a brief discussion of Fichte's arguments for the three stages of freedom, I will examine the question of whether it is possible to attribute a broader notion of liberty to Kant. In view of Fichte's concepts, it is necessary to consider why Kant had defined freedom in such a narrow moral-philosophical way and whether it is legitimate to understand his notion of freedom in a broader sense.

Fichte's Notion of the Three Stages of Freedom in the 1798 *System of Ethics*

In his *System of Ethics* Fichte develops the concept of the three stages of freedom, for which the basic act of "self-positing" of the I is a systematic prerequisite. The simplest manifestation of subjective freedom is spontaneity, which enables conceptual thinking. This first stage of

⁴GMM 99, AK 4:452–3.

formal freedom is followed by the second stage of material freedom. For Fichte, this is the ability to define purposes and intentions for our acts by means of the concepts formed, which reverses the natural and mechanical course of events by a type of causality that is grounded in reason. The final and highest stage can be reached through the moral autonomy of human reason. I now proceed to a more thorough systematic discussion of these stages of freedom.

Formal Freedom

According to Fichte, conceptual thinking liberates our notions of things from immediate sensory perception and intuition. Conceptual thinking allows negation and therefore an idea about things that are not present. Fichte explicitly calls this stage “formal freedom” and sets it as causality of the subject beside the causality of nature. He says:

What ensues from the drive is not something brought about by nature, for the latter is exhausted with the generation of the drive. Instead, this is something I bring about—employing, to be sure, a force that stems from nature, but one that is no longer *nature’s* force but is *mine*, because it has come under the sway of a principle that lies above all nature, under the sway of the concept. Let us call this kind of freedom “*formal freedom*.” Whatever I do with consciousness, I do with this kind of freedom. Someone might therefore follow this natural drive without exception, and yet he would still be free in this sense of the term—so long as he acted with consciousness and not mechanically; for the ultimate ground of his acting would not be his natural drive, but rather his consciousness of this natural drive.⁵

The transparency of the concept and the consciousness of it is what gives man a kind of freedom that the mechanism in nature does not have. At this stage, thinking beings know what they are doing even if they do not intervene purposefully in the course of their acts.

By means of the insights gained through concepts, formal freedom enables us to reflect even on those acts that are motivated by the causality

⁵ SE 129, GA 1/5:129, SW 4:135.

of nature and the influence of the subconscious or of society. But formal freedom is obviously not capable of integrating those reflected insights responsibly and giving the course of events another direction, if necessary. This only becomes possible in the next stage, when the concepts accompanying matters allow the individual an extent of reflection that makes changes imaginable and ultimately desirable. Here the second stage begins, which Fichte calls material freedom. But he does not explain in detail how we can make the transition from merely reflecting on something by means of concepts to acting responsibly.

If we follow Fichte, we can say that reflecting actions by means of concepts just results in a descriptive representation of occurrences, events, or experiences. Even though Fichte did not provide for this with sufficient clarity, it is important to see that reflecting thoroughly about what is represented by concepts enables us to understand the hidden and the obvious values attached to occurrences, events, and experiences that are always subject to evaluation. According to the epistemological approach in the 1794/95 *Grundlage* and the theorem of interest developed in §11 of the *System of Ethics*, a concept is connected to an act of striving and willing that is directed intentionally, which means that conceptual cognition is supplemented by positive or negative judgment.⁶ Through a reflection upon things and their conceptual evaluation, the individual is able to judge these assessments and either to accept or to reject them consciously. To put it in modern terms, it becomes possible in this way to question automatic assessments, which often go unnoticed. This in turn leads to the new assessment and judgment of occurrences, events, and experiences, which can subsequently influence actions.

Fichte does not pay sufficient attention to the fact that the conceptual penetration of facts can also reveal that we are powerless to act differently and are driven by natural impulses. In spite of this, the ability to penetrate facts conceptually offers an opportunity, even in the case of addictive behavior, that the insight gained through concepts imparted with the help of others then leads to better motivation and therefore action. The feeling of being powerless in the face of natural drives going in the wrong

⁶Cf. SE 138, GA 1/5:137, SW 4:145.

direction for whatever reason makes us conscious of a feeling of displeasure that can in the long run awaken and motivate a desire to assert oneself.

This goes to show that formal freedom is at first sight nothing but the consciousness that accompanies things conceptually. To this we can add—going beyond Fichte’s explicit tenets—that conceptual consciousness does not simply accompany things with insights, but can reveal that between the insight and the action there comes a feeling of acceptance, neutrality, or opposition. This feeling can now lead to a higher stage of freedom, if the possibility is used to strive for this higher stage, driving us to make changes and to explore other possibilities.

Material Freedom

A higher level of freedom is reached when the subject is not only free in the sense that it has concepts and language and can accompany insights with clear self-consciousness, but when it—as Fichte says—posits itself as free. By this he means the kind of spontaneity that makes a directed use of intentional volition possible. In this stage of freedom, the subject knows that it is a goal-setting causality; it sees itself as a determining and self-determining being, as an I. Fichte calls this stage material freedom and distinguishes it from formal freedom:

The former [the formal freedom] consists merely in the fact that a new formal principle, a new force, comes upon the scene, without making the slightest change in the material contained in the series of effects. In this case it is no longer nature that acts, but a free being, even though the latter brings about exactly the same thing that nature itself would have brought about if it could act. Freedom in the second sense [viz., material freedom] consists in this: not only does a new force come upon the scene, but there is also a completely new series of actions, with respect to the content of the same. Not only does the intellect engage from now on in efficacious action, but it also accomplishes something completely different from what nature would ever have accomplished.⁷

⁷ SE 132–3, GA 1/5:132, SW 4:139.

Fichte summarizes this again and stresses that in the first formal stage of freedom the free being acts as nature would act if it were subjectivity. Then he goes on to discuss the second stage, which he calls material freedom. It is characterized by conscious, intentional possibilities of action. Every long-term goal makes it necessary to define sub-goals and to reach them by actions. This involves choosing one's overall goals and each sub-goals consciously from several options. Generally there are several strategies for action possible for any one goal of action. Additionally, every goal can be subdivided into many units of action and stages, which in turn offer a broad spectrum of possible actions to choose from as soon as one unit of action has been concluded. According to Fichte, material freedom means deciding between various possibilities that are all open to the willing subject at these points. For Fichte, freedom is not an amorphous projection of self-determination that disregards nature completely. The spectrum of possibilities that can be defined by our insight through reflection is, of course, limited by the horizon of possibilities determined by the ineluctable laws of nature. Making the best use of this horizon of possibilities and reaching the most favorable decision is the result of an interaction between our personal scope for free action and the sound judgment of external possibilities as well as one's own. To be able to do this, one needs a profound knowledge of the various levels that is based on differentiated concepts. Additionally, it is necessary for the self and for the other individuals participating in the project to have the capacity for sufficient and goal-directed motivation to act. When discussing material freedom, it becomes clear in the end that insight and action sometimes converge in this stage and sometimes become opposing internal forces.

Moral Freedom

In Fichte's third stage of freedom, as defined in his *System of Ethics*, the subject puts the ethical law into practice and reaches full moral autonomy. I can only discuss Fichte's definition of morality

as “*freedom—for the sake of freedom*”⁸ very briefly here. This formula contains two concepts of freedom:

In its second occurrence, we are dealing with an objective state that is supposed to be brought about—our ultimate and absolutely final end: complete independence from everything outside of us. In its first occurrence, we are dealing with an instance of acting as such and with no being in the proper sense of the term, with something purely subjective. I am supposed to *act freely* in order to *become free*.⁹

Fichte believes that observing moral laws is mankind’s ultimate purpose. The final end of the moral law, however, “is absolute independence and self-sufficiency, not merely with respect to our will, for the latter is always independent, but also with respect to our entire being.”¹⁰ The independence of the will stated here, which penetrates the entire being of a subject, obviously implies that undivided moral volition is the prerequisite for moral action in the full sense of the word. Fichte points out that in his *Groundwork* Kant had already said “that it is only through the predisposition to morality that a rational being reveals itself as something *in itself*: that is, as something self-sufficient and independent.”¹¹ But in most cases the moral will has the task of defeating those instinctive drives that are in opposition to the moral will. Therefore the concept of freedom has a second meaning. The finite being has to transform itself into an absolutely moral being in order to become identical with itself.

In contrast to Kant, but in a remarkable analogy with Spinoza, Fichte argues that not only do affects and inclinations correlate with an instinctive drive, but cognitive acts like thinking or cognition do so as well. Fichte might owe this systematic conception of instincts and drives to

⁸ SE 145, GA 1/5:143, SW 4:153.

⁹ SE 145, GA 1/5:143, SW 4:153.

¹⁰ SE 198–9, GA 1/5:191, SW 4:209.

¹¹ SE 147, GA 1/5:145, SW 4:155. See also GMM 84–5 and 102, AK 4:435 and 458.

Spinoza¹² or to Karl Leonhard Reinhold.¹³ Duty is tied to what Fichte calls pure drive, whereas the sensuous inclinations that support duty or oppose it and the drives that accompany the inclinations are called “material.” Fichte explains:

As we have now seen, the ethical drive is a mixed drive. It obtains its material, toward which it is directed, from the natural drive; that is to say, the natural drive that is synthetically united and fused with the ethical drive aims at the same action that the ethical drive aims at, at least in part. All that the ethical drive obtains from the pure drive is its form. Like the pure drive, it is absolute; it demands something purely and simply, for no end outside of itself.¹⁴

Fichte emphasizes that the categorical imperative is not a drive as such and therefore not a pure drive, but that it is brought about by reason thanks to reflection and concepts. The pure drive is a causality that is not a causality. That means that the causality of reason cannot be simply equated with causality in nature, which determines the laws of natural appearances.

Like Kant, Fichte also distinguishes the merely exterior accordance of duty and maxims of action that is motivated by fear, obedience, habit, and so on, as one form of legality—which just means obeying moral laws outwardly—from true morality that results from wanting and fulfilling what is requested by duty because of one’s internal moral disposition. We can conceive of many intermittent stages between mere legality, which has to be located at the point of transition from material freedom to moral freedom, and the pure morality of ethical

¹² For the importance of the drive in Spinoza, cf. Thomas Cook, “Der Conatus: Dreh- und Angelpunkt der *Ethik*,” in *Baruch de Spinoza: Ethik*, ed. Michael Hampe and Robert Schnepf (Berlin: Akademie Verlag, 2006), 151–70.

¹³ The seventh letter of Karl Leonhard Reinhold’s second volume of *Briefe über die Kantische Philosophie* (1792) details the problem: “On the hitherto unrecognized differences between the selfish and the unselfish drives, and the two drives and the will (Ueber den bisher verkannten Unterschied zwischen dem eigennützigem und uneigennützigem Triebe, und zwischen diesen beyden Trieben und dem Willen.)” See Karl Leonhard Reinhold, *Briefe über die Kantische Philosophie*, kommentierte Ausgabe, vol. 2, ed. Martin Bondeli (Basel: Schwabe, 2008), 161–82 (in original edition 220–61).

¹⁴ SE 144, GA 1/5:143, SW 4:152.

acts. The formal aspects of moral duties are filled with material content and the maxims formulated in each case owe their existence to judgment and the positing of the I.

When defining what is material in moral acts, Fichte puts a lot of trust in the feeling of approval, that is to say, our conscience. He states, “the absolute criterion for the correctness of our conviction concerning duty, which is what we have been seeking, would therefore be a feeling of truth and certainty.”¹⁵ It is astonishing that Fichte does not have any doubt that this feeling of certainty and hence our conscience might be under an immoral influence, or might have been shaped by one, and therefore might not indicate what is true and morally good. The feeling of certainty shows the identity of the I with itself. A reflection on the nature of this identity shows that Fichte thinks of a consistency of the I with its moral duties.

Ultimately, correct moral actions cannot be found via argumentation, because arguments necessitate endless additional arguments and proofs that can only be silenced by an immediate feeling. Fichte writes:

The feeling of certainty, however, is always an immediate harmony of our consciousness with our original I—nor could things be otherwise in a philosophy that begins with the I. This feeling never deceives us, since, as we have seen, it is present whenever there is complete harmony of our empirical I with the pure I, and the latter is our sole true being, all possible being and all possible truth.¹⁶

In view of this, it is not surprising that Fichte sees one main cause of evil in thinking. In §16, which is dedicated to the question of evil, he stresses that “human nature is originally neither good nor evil. Only through freedom does it become either of these.”¹⁷

In the appendix to §16 of the *System of Ethics* Fichte explains that the source of all evil is our original inertia and states that our “original

¹⁵ SE 159, GA 1/5:156, SW 4:167.

¹⁶ SE 161, GA 1/5:158, SW 4:169.

¹⁷ SE 179, GA 1/5:174, SW 4:188.

laziness or inertia (*Trägheit*) with respect . . . to reflection . . . would be a truly positive radical evil.”¹⁸ The intentional decision to start thinking clearly, that is, to engage in what Fichte calls formal freedom, is the first prerequisite for implementing morality. Fichte emphasizes that inertia is rooted in human nature and that it also affects reflection. Inertia leads to cowardice, conformism, and falsehood, which means that the individual prefers the comfort of saying “yes” to risking the discomfort of opposition.¹⁹ For Fichte, positing the I is an act of active self-positing, which requires enormous active strength. To rise up to the level of reason and hence to the highest form of freedom means to engage in maximum activity and to leave behind all inertia and consequently also the inertia of not wanting to think.

According to Fichte, evil that springs from being the slave of natural instincts and sensuous inclinations is only a minor cause of evil. But this does not mean that a human being, a rational being, is exempt from all guilt and responsibility. Whatever his or her formative social conditions may have been, a human being is a rational being and is therefore able to choose an easier or a more difficult path and to use the forces of thinking, judgment, and reason to determine what should be done:

Despite all the evil examples and all the perverted philosophical arguments, it remains true that a human being ought to raise himself above the laws of nature, and he is also capable of doing this; and it always remains his own fault if he does not do so. For after all, none of these external circumstances exercise any causality upon him; it is not *they* that operate in him and through him, but it is he himself who determines himself in response to a stimulus from the latter.²⁰

The danger that individuals opt for what is evil when it is not the natural drive that governs them, but the drive of self-determination, is far greater. This drive wants to rule, wants to be active, and is based on a kind of

¹⁸ SE 189, GA 1/5:182, SW 4:199.

¹⁹ Cf. SE 191–3, GA 1/5:185–6, SW 4:202–3.

²⁰ SE 175, GA 1/5:171, SW 4:184–5.

lawless arbitrariness which manifests itself in selfish arrogance and a lust for power. Fichte equates this attitude with a desire for sacrifice and heroism which does not strive for moral reason but for unlimited power of the self. This attitude is not rooted in inertia, because it requires much commitment, activity for others. In spite of this, inertia does play a role here, since the individual is incapable of respecting moral laws. It is the arbitrary power of the self and the lawless drive for absolute independence that determines what has to be done,²¹ and the subject is thoughtless and cannot see what the moral law of reason really requests. Thus it is not only inertia, but a kind of hyperactivity in the wrong place that, in Fichte's opinion, makes evil possible. In any case, evil has its roots in human freedom, irrespective of the extent of liberty or the stage of freedom the individual has reached and is able to put into practice.²²

Fichte thinks that freedom is realized step by step and that a certain stage that we have reached is not something that has been gained forever, but is something that happens gradually. The stages serve as descriptions and differentiations and make many other points of transition conceivable and possible.

Kant's Notion of Transcendental Practical Freedom in Relation to Fichte's Three Stages

In the critical period Kant first introduced the notion of transcendental freedom in the famous Third Antinomy in the *Critique of Pure Reason*.²³ For him, this context is so important that he not only refers to it several

²¹ Cf. SE 180–1, GA 1/5:175–6, SW 4:190–1.

²² Cf. SE 173, GA 1/5:169, SW 4:183.

²³ A446/B474, A533–7/B560–5, A802/B830. For a detailed study of Kant's earlier precritical use of the notion of "transcendental freedom," see Heiner Klemme, *Kants Philosophie des Subjekts: Systematische und entwicklungsgeschichtliche Untersuchungen zum Verhältnis von Selbstbewußtsein und Selbsterkenntnis* (Hamburg: Meiner, 1996), 82–95. Dieter Schönecker focuses his interpretation on the precritical and critical difference between transcendental and practical freedom starting with Kant's passage in the Canon chapter of the *Critique of Pure Reason*, A801–4/B829–31, in his book *Kants Begriff Transzendentaler und Praktischer Freiheit* (Berlin: de Gruyter, 2005).

times in his writings about moral philosophy but also presupposes that his readers are familiar with it. In spite of this, he seems to think that the idea of freedom does not really have its place in epistemology, which was discussed in the *Critique of Pure Reason*, but in moral philosophy, because it is one of the three postulates of pure practical reason. In his *Critique of Practical Reason* Kant explains how to conceive the concept of freedom in its strictest sense, which he defines explicitly as the transcendental sense of freedom:

But if no determining ground of the will other than that universal law-giving form can serve as a law for it, such a will must be thought as altogether independent of the natural law of appearances in their relations to one another, namely the law of causality. But such independence is called *freedom* in the strictest, that is, in the transcendental, sense. Therefore, a will for which the mere lawgiving form of a maxim can alone serve as a law is a free will.²⁴

Fichte does not explicitly discuss the distinction between transcendental and nontranscendental freedom that Kant makes here or in other contexts, although Kant's practical philosophy and his theory of freedom were eminently important to him, as he wrote in the aforementioned draft of a letter to Friedrich August Weißhuhn in 1790.²⁵ While Kant links the complete independence of freedom and of the determination of the will to negative freedom, which he discussed earlier in his career in the *Groundwork*, the kind of lawfulness of freedom referred to here is related to positive freedom.²⁶ Kant claims further: "thus freedom and unconditional practical law reciprocally imply each other."²⁷ The strict concept of transcendental freedom in a positive sense focuses on the autonomy of reason, which has to obey the moral law because of its

²⁴ CPpR 162, AK 5:29.

²⁵ Cf. GA 3/1, no. 63.

²⁶ For Kant's conception of autonomy cf. Onora O'Neill, "Autonomy and the Fact of Reason in the *Kritik der Praktischen Vernunft* (§§7–8: 30–41)," in *Immanuel Kant: Kritik der praktischen Vernunft*, ed. Otfried Höffe (Berlin: Akademie Verlag, 2011), 71–85, 79.

²⁷ CPpR 162, AK 5:29.

nature. The content of a moral action is determined by maxims that have to be measured against the requirements of the moral law.

Whereas Fichte uses the feeling of identity with oneself and hence our conscience as a yardstick to determine whether a moral maxim is right or wrong, Kant developed several formulae of the categorical imperative in the *Groundwork* that allow us to examine whether a maxim has been formulated by autonomous reason according to the law of freedom—as Kant calls the moral law—or not.²⁸ Autonomous reason has to be strictly independent of any sensuous influence. I can only point to this briefly here without going into detail.²⁹

In the section entitled “Critical Examination of the Analytic of Pure Practical Reason,” Kant reflects on the relationship between theoretical and practical reason. This leads him to a comparison between causality in nature and the causality of reason. He presupposes that readers know what he said about this point in the *Critique of Pure Reason* in the context of the Third Antinomy of Freedom. But he takes the problem up again to consider it from another point of view. Here he juxtaposes the concept of transcendental freedom with the concept of merely psychological freedom. Psychological freedom means that the moral being is one that acts in space and time and is limited by events in space and time:

They [the determining representations in respect of the subject] are always *determining grounds* of the causality of a being insofar as its existence is determinable in time and therefore under the necessitating conditions of past time, which are thus, when the subject is to act, *no longer within his control* and which may therefore bring with them psychological freedom (if one wants to use this term for a merely internal chain of representations in the soul) but nevertheless natural necessity; and they therefore leave no *transcendental freedom*, which must be thought as independence from everything empirical and so from nature generally, whether it is regarded as an object of inner sense in time only or also of outer sense in both space

²⁸ Cf. GMM 73–87, AK 5:421–38.

²⁹ For a discussion of the different formulae of the categorical imperative, see Thomas Pogge, “The Categorical Imperative,” in *Grundlegung zur Metaphysik der Sitten: Ein kooperativer Kommentar*, ed. Otfried Höffe (Frankfurt: Klostermann, 2003), 172–93.

and time; without this freedom (in the latter and proper sense), which alone is practical a priori, no moral law is possible and no imputation in accordance with it. Just for this reason, all necessity of events in time in accordance with the natural law of causality can be called the *mechanism* of nature, although it is not meant by this that the things which are subject to it must be really material *machines*.

Psychological freedom is also called comparative freedom and strongly distinguished from the transcendental concept of freedom. The latter is absolutely independent of sensuous causes, whereas the comparative freedom has to be seen in accordance with natural causality. Here Kant uses the widely quoted metaphor of a turnspit for comparative freedom. This view now leads us to the critical question of whether formal freedom, which is Fichte's first stage, is covered by Kant's verdict and is basically nothing but psychological or comparative freedom.

Fichte defined *formal* freedom as follows:

Whatever I do with consciousness, I do with this kind of freedom. Someone might therefore follow this natural drive without exception, and yet he would still be free in this sense of the term—so long as he acted with consciousness and not mechanically; for the ultimate ground of his acting would not be his natural drive, but rather his consciousness of this natural drive.³⁰

Here Fichte seems to express exactly the view that Kant attributes to Leibniz, namely that we accomplish a series of actions in time and space driven by an idea or ideas:

Here one looks only to the necessity of the connection of events in a time series as it develops in accordance with natural law, whether the subject in which this development takes place is called *automaton materiale*, when the machinery is driven by matter, or with Leibniz *spirituale*, when it is driven by representations; and if the freedom of our will were none other than the latter (say, psychological and comparative but not also transcendental, i.e., absolute) then

³⁰SE 129, GA 1/5:129, SW 4:135.

it would at bottom be nothing better than the freedom of a turnspit, which, when once it is wound up, also accomplishes its movements of itself.³¹

However, it seems to me that to equate Fichte's formal freedom with an automatic machine driven by ideas or—even worse—with the freedom of a “turnspit” does not do justice to the notion. It is true that not only the effect but also the cause is the same in both cases, because the notion does not serve as a cause in Fichte's concept of formal freedom. Kant obviously claims that transcendental freedom is not a thought alone produced by spontaneity but has to be the causal source of an action. In contrast, Fichte emphasizes the freedom that results from the spontaneity of thinking, which accompanies actions and events through time. In this case, it is not the idea or the concept that determines the event due to the causality of reason. The natural event occurring in time, that is to say, natural causality, is supplemented by another phenomenon: conscious experience. This is, in Fichte's view, an aspect of a prior notion of freedom, because this notion is an intelligible instance itself outside of time. Kant does not consider this aspect of spontaneity as a prior kind of freedom in his discussion of comparative freedom.

When it comes to defending the possibility of a moral cause of sensible phenomena, Kant comes to conclusions that are not dissimilar to those Fichte reaches in his *System of Ethics*. Kant writes:

For, the *sensible life* has, with respect to the *intelligible* consciousness of its existence (consciousness of freedom), the absolute unity of a phenomenon, which, so far as it contains merely appearances of the disposition that the moral law is concerned with (appearances of the character), must be appraised not in accordance with the natural necessity that belongs to it as appearance but in accordance with the absolute spontaneity of freedom. One can therefore grant that if it were possible for us to have such deep insight into a human being's cast of mind, as shown by inner as well as outer actions, that we would know every incentive to action, even the smallest, as well as all the external occasions affecting them, we could calculate a human being's conduct for the future with as much certainty as a lunar or solar eclipse and could nevertheless

³¹ CPrR 217, AK 5:97.

maintain that the human being's conduct is free. If, that is to say, we were capable of another view, namely an intellectual intuition of the same subject (which is certainly not given to us and in place of which we have only the rational concept), then we would become aware that this whole chain of appearances, with respect to all that the moral law is concerned with, depends upon the spontaneity of the subject as a thing in itself, for the determination of which no physical explanation can be given. In default of this intuition, the moral law assures us of this difference between the relation of our actions as appearances to the sensible being of our subject and relation by which this sensible being is itself referred to the intelligible substratum in us. From this perspective, which is natural to our reason though inexplicable, appraisals can be justified which, though made in all consciousness, yet seem at first glance quite contrary to all equity.³²

As Kant makes clear in this passage, he associates freedom seen as the autonomy of reason with the idea of absolute spontaneity. Thus absolute spontaneity has moral-philosophical connotations and manifests itself when the subject exerts freedom on the basis of the autonomy of reason.

Spontaneity is also mentioned frequently in the context of theoretical reason, where Kant would not talk about freedom. To Kant, spontaneity in the determination of concepts and of judgment means that it is something that does not act in space and time and is the opposite of sensibility and receptivity.

When a subject acts with a rational end, it practices what Fichte calls intentional or material freedom. This kind of freedom implies that “not only does a new force come upon the scene, but there is also a completely new series of actions, with respect to the content of the same. Not only does the intellect engage from now on in efficacious action, but it also accomplishes something completely different from what nature would ever have accomplished.”³³

Now one question arises: Is Kant's theory of freedom, which he always wants to restrict to the autonomy of moral reason existing outside space and time—or, to put it differently, constituting a thing in itself not

³² CPtR 219, AK 5:99.

³³ SE 132–3, GA 1/5:132, SW 4:139.

rooted in space and time—is this theory compatible with Fichte’s stages of freedom? Kant writes:

As a rational being, and thus as a being belonging to the intelligible world, the human being can never think of the causality of his own will otherwise than under the idea of freedom; for, independence from the determining causes of the world of sense . . . is freedom. With the idea of freedom the concept of *autonomy* is now inseparably combined, and with the concept of autonomy the universal principle of morality, which in idea is the ground of all actions of *rational beings*, just as the law of nature is the ground of all appearances.³⁴

Kant talks about negative and positive freedom, which are both necessary to enable autonomy as a kind of moral agency independent of the constraints of the senses or drives. In that sense of independence, freedom is negative. But autonomous acting also has to conform to the moral law. Admittedly, this is a merely formal purpose of autonomy, but a positive one in the context of the determination of maxims. A subject that formulates a maxim for its actions realizes a purpose, an intention, and determines what the content, the material, of its moral action will be on the basis of these maxims. But in order to realize an intention, a moral purpose, the subject needs conceptual thinking. Kant emphasizes continually that to realize pure practical reason one has to make use of what pure theoretical reason has brought about in the *Critique of Pure Reason*. Therefore we can say that Fichte has defined stages of freedom that Kant could not or would not accept in this form, since the only authorized notion of freedom for him is the autonomy of pure practical reason in a transcendental respect. In spite of this, the different aspects of freedom that Fichte’s analysis has made explicit are necessarily and implicitly contained in the negative and positive freedom of the autonomous reason. Kant did not dare to postulate pure theoretical aspects of freedom, although the absolute spontaneity of reason is a central moment of his epistemology. He was convinced that there is no proof

³⁴ GMM 99, AK 4:452–3.

of the reality of speculative freedom, whereas the moral law provides objective reality to the notion of practical freedom.³⁵

In this respect it is of interest that Otfried Höffe has proposed including a theory of stages of freedom in Kant's thinking.³⁶ Höffe intends to counter the arguments that brain researchers put forward to refute the assumption that a human being is free. He entitles his essay "Der entlarvte Ruck. Was sagt Kant den Hirnforschern?" ("A Revealing Jolt: What Does Kant Have to Say to Brain Researchers?"). Höffe refers to the didactic example of the "Third Antinomy of Freedom," in which Kant treats the coexistence of determinism on the basis of natural causality and freedom, as a conceivable possibility even though it cannot be proved theoretically. Höffe rightly says that brain researchers observe short-term impulses in their experiments (*Libet*-experiments), whereas in reality decisions are often made over a much longer period of time. According to Höffe, even children, plants, and animals have this kind of elementary freedom and can do this or that, can make decisions. He sees rational, goal-oriented decisions as a higher level. The third and highest level is to him the full autonomy of the will that reveals man as a moral being and makes him act according to the notion of what is good. Whether these forms of freedom—that is, of acting within the horizon of what is morally good, what ought to be done, and what is required of us by duty—can be captured in an experiment is an open question. In his essay Höffe has tried to discover three possible stages of freedom in Kant's theory, just as Fichte defined them in his *System of Ethics*.

With respect to the important theory of different stages in Fichte's notion of freedom, I want to mention that further discussion is needed in order to work out the strange lack of consistency in the concepts he uses. We might be able to show that his idea of reason is modified to generate a control mechanism that turns freedom into a lack of freedom. A type of reason that controls itself absolutely is driven by destructive coercion,

³⁵ A557–8/B585–6; CPpR 178–80, AK 5:47–50.

³⁶ Otfried Höffe, "Der entlarvte Ruck: Was sagt Kant den Hirnforschern?" in *Hirnforschung und Willensfreiheit: Zur Deutung der neuesten Experimente*, ed. Christian Guyer (Frankfurt: Suhrkamp, 2004), 177–82.

which, for instance, puts narrow limits on political freedom. This can be observed in Fichte's *Foundations of Natural Right According to the Principles of the Wissenschaftslehre*, which appeared in 1796/97. As one of the problematic consequences, can we accept that a dangerous criminal loses his human dignity because he should be "an outlaw devoid of rights" who "is declared to be a thing, a piece of livestock," as Fichte suggests?³⁷ In any case, Fichte is of the opinion that the law is only needed as long as morality has not become a binding reality for everyone. Kant, on the contrary, provides a pragmatic solution in his *Perpetual Peace*. For him politics has the purpose of ending all wars, but it is open to conflicts, dissent, different values, and so on. This means that the chances for political freedom are better in Kant's republicanism. But this is an issue that I shall have to address in another paper.³⁸

Violetta L. Waibel is Professor of Philosophy at the University of Vienna. Author of numerous articles, she has published *Hölderlin und Fichte, 1794–1800* (Schöningh, 2000) and also edited (together with D. Breazeale and T. Rockmore) *Fichte and the Phenomenological Tradition* (De Gruyter, 2010). Her main interests are metaphysics, history of philosophy, Kant and German idealism, aesthetics, and romanticism.

³⁷ See §20 "On Penal Legislation": FNR 226–48, SW 3:260–85. For the question of being an outlaw devoid of rights, see FNR 241–2, SW 3:278–9.

³⁸ The present chapter has been translated by Susanne Costa-Krivdic and is a slightly revised version of a paper presented at the conference "Transcendental Philosophy and Metaphysics," held in Osaka, Japan, April 23–25, 2015, which was organized by Halla Kim, Yuko Irie, and Steven Hoeltzel.

4

Fichte, Transcendental Ontology, and the Ethics of Belief

Steven Hoeltzel

Can a transcendental account of the necessary conditions for *knowing* epistemically underwrite, or otherwise rationally require, some specific and accordant understanding of the basic nature of *being*? If so, then supposing that transcendental epistemology is tenable, transcendental ontology may prove to be tenable also. Moreover, if the conceptual and justificatory connections between these two types of positions turn out to be sufficiently tight, then the attempt to articulate a transcendental epistemology that neither presupposes nor privileges any particular ontology might even prove to be unsustainable.

This chapter outlines a reading of Fichte's Jena-era *Wissenschaftslehre* (1794–1801) according to which that system is founded in a transcendental epistemology *and* issues in a transcendental ontology: specifically, the broadly theistic, immaterialist account of autonomously unfolding

S. Hoeltzel (✉)

Department of Philosophy and Religion, James Madison University,
Harrisonburg, Virginia, USA
e-mail: hoeltzsh@jmu.edu

© The Author(s) 2016

H. Kim, S. Hoeltzel (eds.), *Transcendental Inquiry*,
DOI 10.1007/978-3-319-40715-9_4

rational being sketched in Book III of *The Vocation of Man*. Such a reading faces several obstacles. First, among present-day interpreters of Fichte's Jena philosophy, the prevailing view seems to be that, within the methodological framework proper to transcendental philosophy *sensu stricto*, any foray into ontology will be indefensible if not incoherent.¹ Second, even if we grant that some sort of transcendental ontology is in some sense defensible, it remains far from obvious that the Jena *Wissenschaftslehre* itself can accommodate any such enterprise: the texts contain statements that *prima facie* preclude attributing ontological ambitions to Fichte's system. I cannot deal adequately with all of these difficulties here, so I focus mostly on surmounting just the first, mainly methodological obstacle to an ontological interpretation of Fichte's position. Along the way, I briefly raise and respond to some of the more pressing text-based objections to this interpretation, but in the main it remains a separate task (underway elsewhere)² to show more convincingly that the texts support the reconstruction I propose.

This reconstruction maps the conceptual path that leads from (i) the extended transcendental argument that underwrites Fichte's central substantive commitments to (ii) the indicated ontological elaboration upon those systematically central claims. These waypoints are linked, I argue, by a radical reconception of the nature of rationality, and thus of rational justification, that Fichte's transcendental epistemology first secures and which his transcendental ontology draws upon and develops. More specifically, his account of the subjective preconditions for object-directed cognition provides deep transcendental foundations for a strong form of nonepistemic justification; on that basis, his philosophy comes to encompass firm, rationally mandated assent to a specific metaphysical vision of the essential nature and basic makeup of what there really is.

To draw the key connection in a different way: Fichte's transcendental theory has important implications for the ethics of belief, especially

¹ See, for example, Daniel Breazeale, "Jumping the Transcendental Shark: Fichte's 'Argument of Belief' in Book III of *Die Bestimmung des Menschen* and the Transition from the Earlier to the Later *Wissenschaftslehre*," in Fichte's "*Vocation of Man*": *New Interpretive and Critical Essays*, ed. Daniel Breazeale and Tom Rockmore (Albany: SUNY Press, 2013), 199–224.

² See below, notes 5, 14, and 56.

concerning whether we can ever be obligated to accept propositions for which we possess no evidence. On his view, we can be so obligated; the pertinent species of obligation is purely rational (but not narrowly ethical) in origin, and it pertains to only those epistemically intractable propositions, assent to which affirms that the basic metaphysical preconditions for the unconditional validity of reason's highest ideal actually obtain. A fully articulated transcendental idealism must therefore comprise a principled commitment to the extrasubjective reality and all-encompassing totality of a morally purposive, immaterial order of rational being. The broadly theistic immaterialism espoused in *The Vocation of Man* thus reflects a transcendently backed, pure-rational nonevidentialism.³ Because this nonevidentialism derives from what Fichte regards as reason's highest requirement, the ensuing "faith" or "belief" (*Glaube*) does not reflect an irrational fideism. And because he ties the operative requirement to an ultimately hyperethical rational ideal—"the law in question is none other than *the concept of absolute self-sufficiency*"⁴—Fichte's nonevidentialism does not simply recapitulate the practical nonevidentialism of its Kantian prototype.⁵

³ It is worth noting, although I can only address this in passing, that some of transcendental ontology's recent proponents argue that this approach will be conceptually tenable only if it forswears the ontotheological orientation characteristic of, inter alia, the Fichtean position adumbrated above. See especially Markus Gabriel, *Transcendental Ontology: Essays in German Idealism* (London: Bloomsbury, 2011). (Gabriel largely credits this post- and anti-Fichtean position to the later Schelling.) I cannot discuss this in detail here, but I would argue that insofar as this position (i) is based on transcendental reflection upon the basic enabling conditions for truth-apt, object-directed cognition, but (ii) tacitly ties philosophical acceptability exclusively to *epistemic* justifiability, it begs the question against Fichte's transcendently grounded and fully explicit account of what counts as philosophically acceptable and why. Fichte, too, sets out from the standpoint of transcendental reflection, but on that basis he develops an account of the (so to say) "transcendentally necessary" nature of finite rationality, and therefore of rational justification; and in the light of that account, transcendental philosophy provides us with irrevocable rational grounds, *nonepistemic* in nature, for assent to an epistemically indefeasible ontotheological worldview. The remainder of this chapter should serve to clarify and substantiate that claim, although I shall not explicitly return to the issue of ontotheology.

⁴ SE 61, SW 4:60. Cf. SE 58, SW 4:56.

⁵ For some discussion of the parallels with Kant's postulates of pure practical reason, see Steven Hoeltzel, "Non-Epistemic Justification and Practical Postulation in Fichte," in *Fichte and Transcendental Philosophy*, ed. Tom Rockmore and Daniel Breazeale (Basingstoke, UK: Palgrave Macmillan, 2014), 293–313.

Transcendental Arguments and Ontological Commitment

The Jena *Wissenschaftslehre*'s systematically central substantive commitments (or so-called "first principles") are first secured via a regressive transcendental argument isolating the necessary enablers of the elementary abstract organization of all truth-apt, object-directed cognition. According to Reinhold's agenda-setting delineation of this abstract structure, "in all consciousness, the subject distinguishes the representation from both the subject and the object and relates it to them both."⁶ On Fichte's view, empirical cognition always comprises this basic framework of subject-object differentiation-relation, but this structure itself stands in need of a transcendental explanation, one that will isolate and describe the several subjective operations whose unified result Reinhold's principle recounts.⁷ These subjective operations are the elementary a priori accomplishments of rationality, understood as a capacity to originate and instate pure organizing forms, which first effect empirical cognition's abstract articulation: "the subject and object do indeed have to be thought of as preceding representation, but not in consciousness qua an empirical mental state, which is all that Reinhold is speaking of. The absolute subject, the I, is not given by empirical intuition; it is, instead, posited by intellectual intuition. And the absolute object, the not-I, is that which is posited in opposition to the I."⁸

Consciousness per Reinhold's principle comprises *objective reference*: the subject's ascription of qualitative data (sensory or affective givens) to some putatively extrasubjective bearer or source. It is therefore, Fichte argues, a necessary condition for the possibility of experience that "the I absolutely counterposits a not-I."⁹ The empirically given qualitative

⁶ Karl Leonhard Reinhold, *Beyträge zur Berichtigung bisheriger Missverständnisse der Philosophen*, vol. 1 (Jena: Mauke, 1790), 267.

⁷ EPW 63–4, SW 1:8. For further context, see Daniel Breazeale, *Thinking Through the Wissenschaftslehre: Themes from Fichte's Early Philosophy* (Oxford: Oxford University Press, 2013), ch. 2.

⁸ EPW 65, SW 1:9–10.

⁹ SW 1:104, my translation.

contents of consciousness, considered strictly as such, constitute no answer to (and cannot, on their own, even motivate) the question as to whether they reflect the existence of anything other than the subject (the I), *its* basic nature, and *its* determinate state. Consequently, the existence of any extrasubjective entity is something that has ultimately to be *posited* (freely countenanced, actively marked out) by the subject: no such state of affairs is or can be simply *given*.¹⁰ The ascription of certain of consciousness's contents to some extrasubjective source must therefore draw upon and deploy a pure, nonsensory notion—that of the not-I as such—whose transcendental employment in categorizing consciousness's contents is the a priori instatement of one basic element in experience's organizing structure.

This is not, however, the transcendently most basic such activity. That status belongs to *self*-positing, in which “the I originally absolutely posits its own existence.”¹¹ Here, the purity and autonomy of properly rational activity, which the subject of such activity apprehends in a direct nonsensory manner, is marked out and affirmed as proper to the rational subject, such that the I can be said to be “posited by intellectual intuition.”¹² The operative nonsensory notion of the I—the simplest and purest of rationality's self-originated ordering forms—then becomes the categorial cornerstone for the further articulation of experience's organizing structure. So, for example, this act of self-positing—the nonsensory singling-out of the subject qua purely and simply rational—enables and informs the transcendently posterior positing of a not-I: a not-I is posited only if the I finds itself faced with qualitative data that are recalcitrant and opaque relative to the autonomy and purity of the rational activity it has posited as proper to itself.¹³

There is a great deal more to the transcendental story, of course, but perhaps the above makes it passably clear that Fichte's account of the

¹⁰ See especially SK 105, SW 1:104.

¹¹ SW 1:98, my translation.

¹² EPW 65; SW 1:10. The direct self-apprehension involved here is “nonsensory” in the sense that what it apprehends is not an opaque sensory given or unchosen affective goad, but self-transparently pure and autonomous intellectual activity.

¹³ See, for example, SK 189, SW 1:209; cf. SE 81–2, SW 4:81–2.

enabling conditions for object-directed cognition opens up upon a basic ontology of rational being qua rational: rational being as the *self-transparent and autonomous origination and instatement of pure organizing form*.¹⁴ Below, I complicate the transcendental account and elaborate upon the indicated ontology, but first it should be stressed that, on Fichte's account, it is a necessary condition for the existence of object-directed cognition *that there exist* transcendental *activities* of the kind just defined. Granted, these are specifically subjective activities, integral to representation, but by itself this is no reason for refusing to countenance their existence altogether, as if a world containing only such activities would be a world "really" containing nothing at all. Nevertheless, Fichte's system is often thought to occupy a standpoint from which claims about what there really is, as opposed to claims concerning how we necessarily represent things to be, are either metaphilosophically prohibited or methodologically put out of reach. I cannot deal adequately with all of the associated difficulties here, but it should be useful to address briefly a few of them.

To begin with the metaphilosophical issue. Fichte himself positions his system with reference to a sharp distinction between the standpoint proper to philosophy and that characteristic of ordinary life. From the latter standpoint, which is our cognitive default, we unreflectively regard sensible things as objectively real causes of our experiences. We take up the philosophical stance, by contrast, only by stepping back from that default standpoint altogether and intellectually suspending its ontological commitments, so as to see it purely and simply as *a way of representing things*, one whose ultimate grounds and overall significance we now seek to understand.¹⁵ Fichte then conducts that inquiry in an idealistic as opposed to "dogmatic" fashion. That is to say, he eschews the naturalistic enterprise of causal explanation, which "tries to explain representations . . . on the basis of an efficacious action

¹⁴ For more on this broadly Kantian conception of the nature of rationality, and on the distinctly Fichtean transcendental argument that seeks to secure it, see Steven Hoeltzel, "The Unity of Reason in Kant and Fichte," in *Kant, Fichte, and the Legacy of Transcendental Idealism*, ed. Halla Kim and Steven Hoeltzel (Lanham, MD: Lexington Books, 2014), 129–52.

¹⁵ See, for example, EPW 435, GA 3/3, no. 440.

of the thing in itself.”¹⁶ Instead, he works out a transcendental account guided by the programmatic postulate that “everything which occurs in our mind can be completely explained and comprehended on the basis of the mind itself.”¹⁷ According to the resulting theory, all seemingly mind-independent reality thus “*originates for us* only insofar as *one does not engage in philosophizing*,” and “as soon as one lifts oneself to the level of speculation . . . *this reality necessarily disappears*.”¹⁸ First we bracket our everyday ontological commitments, and then we transcendently dissolve everything of that sort into some manner of subjective representing. Thus it begins to appear that, from Fichte’s position, any *philosophical* claims about what there really is would represent a relapse into the sort of belief in an extrasubjective reality that good philosophy brackets and explains away—so that, for all philosophically legitimate purposes, “*being and being posited* . . . must be one and the same.”¹⁹

In spite of appearances, I think, such statements do not in fact disavow the intention to develop a philosophy that will truly characterize what there really is. Instead they signal the philosophical rejection of just one type of ontology, one that Fichte’s philosophical project initially brackets and which his transcendental theory subsequently undermines. This is, of course, the ontology tied to and an ingredient in our unreflective everyday standpoint, which tacitly takes *thinghood*—being, implicitly understood *as* enduring impassive subsistence requiring no

¹⁶ IWL 20, SW 1:435. Cf. IWL 20, SW 1:436: “Dogmatism wishes to use the principle of causality to explain the general nature of the intellect as such.”

¹⁷ EPW 69, SW 1:15. Fichte’s procedure here is not as question-begging as it might seem. He views the idealism-dogmatism dispute as epistemically undecidable at the outset (that is, in advance of the actual construction of theories of both kinds), so he initially opts for the position that best aligns with his own prephilosophical convictions. But he also argues that, once a thoroughgoing idealism has actually been articulated, it will prove superior to any possible dogmatism on grounds of simplicity and explanatory power. (See especially IWL 15–25, SW 1:429–40.) Thus the idealistic principles assumed at the outset are thought to be epistemically vindicated at the end of inquiry. For further discussion, see Breazeale, *Thinking Through the Wissenschaftslehre*, ch. 11.

¹⁸ EPW 434, GA 3/3, no. 440.

¹⁹ SK 172, SW 1:188. See also SE 23, SW 4:17: “There is being only for an intellect.”

contribution from consciousness²⁰—to be something *real*. In Fichte’s ontologically unfortunate terminology, “being” refers specifically to this thing-like *type* of being, not to existence or reality *überhaupt*. “Being” in this specified sense is explicitly opposed and philosophically subordinated to *acting*,²¹ understood as ongoing self-initiated self-actualization.²² Thus the “abstraction from all being” whereby one ascends to Fichte’s philosophical standpoint²³ by no means entails a total disengagement from the issue of what there really is. The required abstraction is rather the principled suspension of the ontological commitments integral to our prephilosophical standpoint, so that these very commitments and the entire outlook that they configure can be accounted for via a higher-order transcendental theory of acting:

The sole thing to which the person who undertakes this act of abstraction continues to cling and proposes to employ as the basis for explaining everything that has to be explained is the conscious subject. Consequently, he must grasp this conscious subject entirely apart from any representation of being, for only in this way will he then be able to show that this subject contains within itself the ground of all being—“being for this subject,” as goes without saying. But if we abstract from all being of and for this conscious subject, then nothing pertains to it but acting.²⁴

The “reality” that “disappears” when we adopt Fichte’s philosophical standpoint is therefore not the reality *epistemologically* affiliated with the aspiration to understand correctly what there really is regardless of what we might mistakenly take there to be. Reality *in that sense* remains a principal object of philosophical interest. The “reality” that vanishes from Fichte’s position is merely the empirically manifest thing-like subsistence

²⁰ See, for example, FNR 27f., SW 3:28.

²¹ IWL 84; SW 1:499.

²² See, for example, FNR 27f., SW 3:28.

²³ IWL 40; SW 1:457.

²⁴ IWL 40, SW 1:457. Cf. IWL 26, SW 1:440: “Idealism considers the intellect to be a kind of *doing* and absolutely nothing more.”

(as opposed to the transcendently implicated rational self-actualization) that we ordinarily—and, philosophically speaking, unjustifiably—ascribe to what there really is. “What exists for the philosopher is acting,” Fichte says, “and acting is all that exists for him; for, as a philosopher, he thinks in an idealistic manner”²⁵—first methodologically suspending and then transcendently undermining the belief in “being” that is basic to our prephilosophical outlook. The result: a philosophy “acquainted with something that is even higher than any being.”²⁶

Fichte holds, I suggest, that *there really exist* subjective activities, rational in nature, that transcendently generate the representations of extrasubjective objects to which we unreflectively ascribe real being from our prephilosophical standpoint. As the foregoing shows, one can make such claims without abandoning the standpoint of Fichte’s philosophy—provided, of course, that one confines these claims to that standpoint, neither seeking to overhaul our unreflective outlook in terms of them,²⁷ nor allowing our accustomed everyday realism, which thinks in terms of “being,” to cloud our higher-order understanding of the transcendental field of rational acting. This restriction of transcendental claims to the philosophical standpoint does not somehow subjectivize those claims or otherwise weaken their import. Rather, it simply preserves the philosophy-founding distinction between our first-order, unreflective outlook on things and this higher-order inquiry into that entire outlook’s basis and basic nature.

Granted, such an inquiry is, relative to our precritical default, entirely optional. No considerations internal to our prephilosophical outlook put us under any epistemic pressure to take up philosophy’s cognitive project. But this fact by itself does nothing to diminish philosophy’s cognitive credentials, and Fichte clearly regards philosophy as the cognitively privileged partner in this relationship. Philosophy poses a basic

²⁵ IWL 84, SW 1:498.

²⁶ IWL 137; GA 1/4:466.

²⁷ “Speculation exists in a totally different world, and anything which is to have an influence upon life must proceed from life. Speculation is only a means for *gaining knowledge* about life” (EPW 434, GA 3/3, no. 440).

and legitimate question concerning the standpoint of life as such, a question that cannot even be formulated from within that standpoint, and philosophy's answer to that question provides a new and higher insight into the origins and overall significance of that whole way of representing reality. "The standpoint of life," Fichte therefore says, "is comprehensible only from the standpoint of speculation."²⁸ Accordingly, a truly well-founded understanding of what there really is can be arrived at only philosophically, that is, "only insofar as . . . the standpoint which transcends life exists alongside the standpoint of life itself . . . One cannot have any knowledge of life without engaging in speculation."²⁹

Note also that the existence claims that concern us here do not involve the philosopher in some sort of "dogmatism" antithetical to Fichte's professed idealism. Dogmatism, for Fichte, is the form of philosophy that would causally yoke subjectivity's existence and operations to some ontologically prior bearer of "being" that just mindlessly, aimlessly, is—the thing in itself, in the sense relevant here. Such a position thus "treats the I merely as a product of things."³⁰ But as we have seen, the transcendental philosopher can make existence claims that neither support nor presuppose any such metaphysical picture. Indeed, were he or she to further maintain that *what there really is* just *is* rational activity, "all the way down," that would be the exact opposite of dogmatism.³¹ And that claim, insofar as it belongs to the higher-order comprehension and clarification of our ordinary way of representing reality, would not overstep the standpoint proper to philosophy.

²⁸ IWL 38 n., SW 1:210 n.

²⁹ EPW 435, GA 3/3, no. 440.

³⁰ IWL 16, SW 1:431.

³¹ Some commentators read Fichte's idealism–dogmatism distinction as marking off a non-naturalistic position that can plausibly source normativity from a naturalism that cannot. See especially Wayne Martin, *Idealism and Objectivity: Understanding Fichte's Jena Project* (Redwood City, CA: Stanford University Press, 1997). I think that there is much to this, but that for Fichte this distinction supervenes upon the more substantially metaphysical distinction drawn above. The discussion to come should make my reasons for saying this reasonably clear.

Pure Reason, Nonepistemic Justification, and Metaphysical Commitment

If the above is correct, then, subject to the restrictions just discussed, Fichte's transcendental idealism is neither metaphilosophically nor methodologically barred from making claims about what there really is. Still, given this idealism's basic explanatory logic, we might suspect that the scope of this enterprise must be so severely circumscribed that it can legitimately issue only in a sort of solipsism. For on Fichte's account, as we saw above, experience's organizing structure of the subject–object differentiation–relation—the pure categorial framework upon which our prephilosophical belief in the real being of *things* depends—is not empirically given but transcendently originated and articulated via pure rational activity. The concrete content that the pertinent rational activities abstractly elaborate upon is supplied by the subject's own recalcitrant sensory and affective states. And if Fichte's theory is to steer clear of dogmatism, then it cannot causally couple these states to any extrasubjective “things in themselves” that would mechanically affect the I. Instead Fichte must make it clear “*that* and *how* the I can evolve, entirely from itself, whatever is to occur therein, without ever emerging from itself or breaking out of its own circle.”³² Given this framework, however, it seems that the philosopher can in principle obtain no evidence backing any belief in the existence of anything other than her own subjectivity, transcendently reconceived. For in order to give an exhaustive account of the transcendently necessary and sufficient conditions for any of her experiences, she need not affirm the existence of anything other than the strictly subjective states and activities that she understands to undergird those experiences transcendently.

As a consequence, Fichte concedes, “with our explanation of consciousness we can never arrive at things that exist independently of us.”³³ Yet this turns out to be nowhere near the end of the ontological story, for he goes on to avow that “the whole sensible world is . . . a mere

³² SK 255, SW 1:289–90.

³³ AD 99, GA 1/5:423.

reflection in mortal eyes of the nonsensible, which alone exists.”³⁴ The latter, he says, which “alone gives meaning, purpose, and value” to sensory representations,³⁵ is a “realm of spirits” or “world of reason”—that is, a “realm of freedom and rational self-activity.”³⁶ All told, “only reason is; infinite reason in itself, and finite reason in it and through it. Only in our minds does it create a world.”³⁷ According to the Jena *Wissenschaftslehre*, then, there is a reality transcendent to human representations, but it is not the sensible world of material nature; it is an immaterial order with an essentially teleological configuration and a broadly theological foundation.³⁸ “I just do not accept,” Fichte says, “the self-sufficient existence of a sensible world”—but this rejection reflects the view of “an *acosmicist*” who is not at all “an *atheist*.”³⁹

But now the question is: How are such claims to be properly justified, given the epistemic intractability (on Fichte’s own epistemology) of the issues they address? Here it is necessary to look past the moralistic and voluntaristic rhetoric of Fichte’s popular presentations in order to grasp the deeper logic of his position. On this account, issues that are epistemically intractable need not be rationally undecidable, because *there is an essentially hyperepistemic ideal, unconditionally valid for rationality as such (because transcendently integral to rationality as such), that nonepistemically justifies some basic cognitive commitments, of a distinctly philosophical kind, pertaining to certain epistemically intractable issues of (rationally) ultimate concern.* The philosopher’s discernment of the way in which this supreme ideal superintends the entire space of reasons will then play two crucial and closely related roles in the development and defense of an affiliated transcendental ontology. First and foremost, it will mandate assent to a special set of descriptive propositions involving some quite specific ontological

³⁴ VM 114, SW 2:308. Cf. VM 109, SW 2:300.

³⁵ VM 99, SW 2:289.

³⁶ VM 94, 95, SW 2:282, 283.

³⁷ VM 111, SW 2:303.

³⁸ Cf. VM 110, SW 2:302–3.

³⁹ AD 180, GA 1/6:54.

commitments. Second, the nature and justificatory force of that mandate will insulate the indicated ontology from criticisms that target this ontology's epistemic insecurity.

But what is that all-important hyperepistemic ideal? How does Fichte make his case for its authority? And how is that supposed to support a series of separate ontological commitments? I can only crudely sketch most of this here, but the essentials are these. According to Fichte, it is "the goal of our entire existence and all of our acting—a goal that is indeed never to be reached but is to be unceasingly promoted—that rational being become absolutely and entirely free, self-sufficient, and independent of everything that is not itself reason."⁴⁰ He holds, as we saw above, that it is a necessary condition of the possibility of object-directed cognition that "the I originally absolutely posits its own existence"⁴¹—that is, that the purity and autonomy of properly rational activity is transcendently marked out and affirmed as proper to the rational subject, in distinction to the opacity and fixity of the empirical contents of its consciousness. Now, note also that if such self-positing serves only *descriptively* to differentiate pure rational activity from recalcitrant qualitative data, then (i) the only distinction thereby affirmed is one internal to that subject's consciousness, and more importantly (ii) the *motivation* is lacking to propel the process onward to the transcendently posterior positing of a *not-I*, figured as an extrasubjective bearer or ground of the unbidden empirical manifestation. "Without a striving," as Fichte would have it, "no object at all is possible."⁴²

It must be, then, that self-positing, in addition to descriptively differentiating rational activity from empirical manifestation, is also such as prescriptively to motivate rational activity in relation to such manifestation. That is, the self-positing subject must somehow and simultaneously be *self-appointed* to sustain the ongoing origination of and instatement of pure ordering forms. In that case, it must be that self-positing marks out and affirms the purity and autonomy of rational activity not only as

⁴⁰ AD 101, GA 1/5:426.

⁴¹ SW 1:98, my translation.

⁴² SK 233, SW 1:264.

essentially constitutive of the transcendental subject but also as *unconditionally normative* for it. Finite rationality thus must have its own non-arbitrary, self-legislated, ultimate end—unconditioned rational activity—because object-directed cognition can eventuate only on the condition that this ideal is authored and approximated to (commitment to it autonomously adopted and enacted) in the transcendental background.

Specifically *epistemic* endeavor approximates to that ideal by assigning rationally unbidden givens (sensory and affective data) to an intelligible order of reason's own design, thereby removing occlusions to rational insight. But specifically *ethical* agency also approximates to that ideal, in this case by evaluating unbidden incitements (given inclinations) according to pure, self-wrought standards, thereby surmounting obstacles to rational autonomy. Pure reason's superordinate ideal itself is thus neither specifically epistemic nor narrowly ethical. It is, so to say, *hyperepistemic* and *hyperethical*, being the absolute optimum of pure rational *being* as such: the origination and instatement of pure ordering forms, unqualified by anything not authored through its own activity, and thus operating independently of any merely given, rationally unbidden limitations. "The whole final purpose of reason," Fichte says, "is its own pure activity . . . i.e., independence from everything which is not itself reason, absolutely unconditioned being."⁴³ Accordingly, on the scale of values vindicated by transcendental idealism, our narrowly epistemic and ethical aims and accomplishments are important indeed—but important *because* a superordinate hyperepistemic and hyperethical ideal of unconditioned rational activity is the ultimate goal of rational being as such.⁴⁴

Now, if this ideal is not merely to describe the highest measure of rational achievement but, beyond that, categorically *and validly* to prescribe the finite rational being's unceasing endeavor toward the indicated optimum, then—supposing, as Fichte does, that "ought" implies "can"⁴⁵—finite rational being cannot take place within just

⁴³ VM 99, SW 2:288.

⁴⁴ For a more detailed treatment of this account of reason and the related transcendental argumentation, see Hoeltzel, "The Unity of Reason."

⁴⁵ See especially IWL 148–9, SW 5:183–4.

any scheme of things. Rather, what there is, and how it is, must be such as to make metaphysically possible the finite rational being's actual approximation to the indicated optimum.⁴⁶ In other words, on Fichte's view *there are metaphysically necessary conditions for the categorically binding authority of the highest rational ideal*—that ideal, commitment to which is a transcendently necessary condition for the possibility of truth-apt, object-directed cognition.

As we saw above, Fichte's transcendental epistemology entails that questions concerning the existence or nature of any extrasubjective reality are epistemically intractable. Thus, no beliefs about such metaphysical matters are *epistemically* justifiable on Fichte's account. Nevertheless, as we can now see, on that account some such beliefs still are *rationally* justified, and indefeasibly so, along roughly the following lines. (1) Truth-apt, object-directed cognition transcendently depends upon tacit commitment to pure reason's hyperepistemic ideal. (2) That commitment constitutes a hyperepistemic requirement to approximate to that ideal. (3) To be subject to that hyperepistemic rational requirement is to possess *nonepistemic* rational grounds for assent to precisely those propositions that must actually be true if approximation to reason's highest ideal is really (that is, metaphysically) to be possible. (4) Belief that the indicated metaphysical conditions actually obtain is thus nonepistemically justified, despite being epistemically unjustifiable—and such belief is also, on the same transcendental grounds, (a) nonepistemically justified *by reason* and (b) epistemically *indefeasible*.

Before we consider some possible problems for this approach to justifying metaphysical claims, let us briefly survey Fichte's account of the indicated metaphysical conditions. First, he maintains, if the subject's states and activities are to have real significance and actual efficacy, as the ideal of rational being demands, then there must exist an extrasubjective world that subjectivity can reflect and affect.⁴⁷ Second, if the subject's endeavors are to be both self-initiated and

⁴⁶ See, for example, IWL 173, SW 5:393.

⁴⁷ VM 69, SW 2:250–1.

actually impactful, as real approximation to the ideal of rational being requires, then what there really is must be immaterially constituted and normatively ordered, not materially constituted and governed by mindless laws.⁴⁸ And third, if it is to be assured that intellectual and ethical activity will always count toward or contribute to something of ultimate value, as rational assent to the ideal of rational being requires,⁴⁹ then there must exist a rationally purposive superintending of the world order, whereby such results are ensured⁵⁰—an “active ordering (*ordo ordinans*)”⁵¹ of which the ethically ordered, immaterial “world of reason” is the ongoing unfolding.⁵² On Fichte’s account, “this living and efficaciously acting moral order is itself God,”⁵³ albeit not in the personified guise of the traditional deity⁵⁴ but instead as “the absolutely pure form of reason”⁵⁵—that is, an autonomously unfolding pure ordering, unconditioned by anything not authored through its own self-transparent self-activity.⁵⁶

Note that the resulting picture of existence is the antithesis of “dogmatism”: it does not figure the finite subject “merely as a product of things, i.e., as an accidental feature of the world,”⁵⁷ but instead sees our subjectivity as purposively engendered by an unconditioned principle that provides for free approximation to a self-legislated ideal. This is also a theological picture, not in the sense that it would causally

⁴⁸ VM 92–3, SW 2:280–1. Cf. IWL 171–4, SW 5:392–4.

⁴⁹ VM 92–3, SW 2:280–1.

⁵⁰ IWL 149, SW 5:184. Cf. VM 95, SW 2:284.

⁵¹ IWL 161, SW 5:382.

⁵² Cf. IWL 180, GA 1/6: 413.

⁵³ IWL 151, SW 5:186.

⁵⁴ IWL 152, SW 5:187. Cf. AD 178, GA 1/6:51.

⁵⁵ SE 143, SW 4:151.

⁵⁶ For more detailed discussion of the indicated ontology and especially its connection to Fichte’s first principles, see Steven Hoeltzel, “Transcendental Idealism and Theistic Commitment in Fichte,” in *The Palgrave Handbook of German Idealism*, ed. Matthew C. Altman (Basingstoke, UK: Palgrave Macmillan, 2014), 364–85.

⁵⁷ IWL 16, SW 1:431.

credit all finite things to the agency of an individuated consciousness, but in the broader sense that it ontologically anchors all existence in a self-active and rationally purposive source that is providential in relation to the highest interests and ultimate aims of the finite rational existences into which it autonomously unfolds.⁵⁸ And all of these claims are of course affirmed only from and strictly for the sake of the standpoint of philosophy: they do not add to or obtrude upon anything that we would take to be true from the unreflective standpoint of everyday life. The indicated commitments are installed at the philosophical level and operative only within and for the purposes of the affiliated outlook.⁵⁹ Still, they do provide our unreflective, pre-philosophical standpoint, as such and as a whole, with a kind of comprehensive clarification concerning its ultimate grounds and overall significance. For ordinary experience can now be philosophically understood to symbolize our situation within, and to elicit and indicate our autonomous interaction with, a supersensible order of being transcendent to our own subjectivity.⁶⁰ Thus, despite experience's *epistemic* disconnect from the real nature of what there is, experience *ethically* regarded remains essentially veridical, insofar as it acquaints us, albeit obliquely, with our being appointed to realize ourselves freely through real and meaningful engagement with a morally ordered ultimate reality.⁶¹

⁵⁸ See VM 109–11, SW 2:300–3. It seems undeniable that assent to this outlook entails the moderation of some of Fichte's very strong early statements as to the character of his idealism. For example: "the intellect cannot be anything passive, because, according to the postulate of idealism, it is what is primary and highest and is thus preceded by nothing that could account for its passivity" (IWL 25, SW 1:440). Still, because on this account what truly exists is rational activity all the way down, with no mindless basis nor any thing-like being, the idealistic and antidogmatic spirit of the position is sustained. Note also that on this account the occurrence and qualitative determinacy of the *Anstoß* and/or *Aufforderung*, while still transcendently inexplicable (IWL 75, SW 1:489–90), no longer stand out as systematically unassimilated brute facts, but instead are supplied with a (very general) teleological explanation.

⁵⁹ See for example Fichte's April 22, 1799 letter to Reinhold and the appended "Fragment": EPW 428–35, GA 3/3, no. 440.

⁶⁰ VM 111, SW 2:303. Cf. VM 114–15, SW 2:308.

⁶¹ See especially IWL 149–50, SW 5:184–5.

Transcendental Philosophy, Pure Practical Reason, and the Ethics of Belief

Let us now consider the logic of this Fichtean position from a somewhat different angle.⁶² For purposes of this discussion, *epistemic grounds* for a proposition should be understood as grounds for assent which are of a sort that any similarly situated rational inquirer would recognize to be reliably indicative of the state of affairs said to obtain by the proposition. Accordingly, *sufficient* epistemic grounds will provide sufficiently strong indications of a proposition's truth to license confident assent.⁶³ Understand *evidentialism* as the view that it is rational for one to assent to a given proposition if and only if one has sufficient epistemic grounds for doing so. And think of the associated *agnostic imperative* as the directive to withhold assent from any proposition for which one lacks such grounds. Finally, let "*rational*" mean "compliant with reason's basic requirements." With these definitions in mind, we can mark out some important consequences of Fichte's position.

First, given Fichte's transcendental epistemology, evidentialism is false. In this connection, that epistemology has three especially apposite entailments. (1) Pure reason "is absolutely practical" or "sets itself an end purely and simply by itself."⁶⁴ (2) This end is not any narrowly epistemic end (justified true belief, absolute knowledge, or the like), but unconditioned rational activity as such (that is, unqualified efficacy in the origination and instatement of pure ordering forms), to which specifically epistemic endeavor is only one avenue of approximation (properly principled volition being another). And (3) questions concerning the existence or nature

⁶² This section's account of Fichte's position is my own, but my understanding of various pertinent issues in and around the ethics of belief, as well as some of the formulations employed here, are indebted to Andrew Chignell, "The Ethics of Belief," in *The Stanford Encyclopedia of Philosophy* (Spring 2016 edition), ed. Edward N. Zalta, <http://plato.stanford.edu/archives/spr2016/entries/ethics-belief/>; and Jeff Jordan, "Pragmatic Arguments and Belief in God," in *The Stanford Encyclopedia of Philosophy* (Winter 2014 edition), ed. Edward N. Zalta, <http://plato.stanford.edu/archives/win2014/entries/pragmatic-belief-god/>.

⁶³ This and the preceding formulation are indebted to (but adapt) Andrew Chignell, "Belief in Kant," *Philosophical Review* 116, no. 3 (2007): 326–7.

⁶⁴ SE 59, SW 4:57.

of any extrasubjective reality are, in principle, epistemically intractable. In consequence of these claims, there will be a special class of specifically philosophical propositions—namely, those that affirm the metaphysical preconditions of the objective validity or binding authority of pure reason’s absolute ideal and the related requirements—assent to which is nonepistemically enjoined by reason despite the total absence of supporting epistemic grounds. Therefore, evidentialism to the contrary notwithstanding, assent to these (epistemically indefeasible) propositions is absolutely rational, qua compliant with reason’s *most basic* requirement. Rejecting any one of these propositions is tantamount to renouncing that very requirement, by denying that some necessary condition for its objective validity actually obtains.

Moreover, on Fichte’s account, simply to suspend judgment on these metaphysical propositions, even on grounds of their glaring epistemic insufficiency, is to refuse to comply with reason’s most basic (and not narrowly epistemic) requirement. More specifically, to do this is (i) to withhold assents that are rationally enjoined by that requirement, (ii) to do so despite the fact that these assents concern questions about which the evidence will always be silent, and (iii) to do so in service to a narrowly epistemic directive: “assent only if given good evidence.” For Fichte, however, the authority of any valid epistemic norm can only derive from its conforming and conducing to reason’s ultimate, underived, hyperepistemic ideal.⁶⁵ And the content of the specific epistemic directive in question here—the agnostic imperative—places it in conflict with reason’s overarching, ultimately overepistemic end, at least on this philosophically decisive point. Accordingly, the agnostic imperative is, for philosophical purposes and at this conceptual juncture, a substantively irrational mandate.

It might now be useful to approach these issues from yet another direction, by presenting an objection to this position and considering Fichte’s response. The objection: Fichte fails to ground adequately his

⁶⁵ See, for example, IWL 148–9, SW 5:183–4—understanding these popular phrasings in terms of the more rarefied transcendental position I have outlined. See also VM 101, SW 2:291 and (especially) IWL 39–40, SW 1:456–7.

claims concerning the objective validity or binding authority of reason's ultimate requirement and, in consequence, cannot convincingly reject evidentialism, flout the agnostic imperative, or elaborate an ontology tailored to reason's ultimate ends. To counter this objection it will not suffice to re-inscribe, *mutatis mutandis*, the points made just above. Those points may obviate tacitly evidentialist or agnostic objections to Fichte's metaphysical elaborations on his transcendental position, but they do so by deploying a notion of nonepistemic justification that is tied to his conception of pure reason as "absolutely practical." The objection now under consideration is that, even if Fichte's transcendental argument in support of that conception goes through, this does not actually suffice to show that the associated hyperepistemic rational requirements really are objectively and categorically prescriptive. And absent of some such demonstration, it is not clear that the affiliated mode of nonepistemic "justification" can provide much (if any) *objectively valid* justification for assent to the kinds of claims under consideration here.

To state more fully the objection: suppose that Fichte succeeds in showing, along the lines laid out above, that commitment to and compliance with a certain pure-rational requirement is a necessary condition for the possibility of experience. (So, please note, the objection grants that pure reason is "absolutely practical," in Fichte's sense, and then presses the following point.) It does not necessarily follow that the requirement thus complied with really is valid, binding, or authoritative, full stop—that is, considered independently of the role that some sort of assent to it plays in the transcendental articulation of the categorial structure required for objective reference and so forth. That we *must already have let ourselves* be bound by that requirement under the specified conditions is not obviously a good reason for thinking that we *ought to be* bound by that requirement come what may. And if we can find no rationally respectable grounds for judging that that requirement really is objectively and categorically prescriptive—and how could we, if, like Fichte, we favor an epistemology according to which extrasubjective reality is always epistemically out of reach?—then it seems that we should judge the ontology that he elaborates in service to this requirement to be rationally groundless in consequence. Rather than simply

faulting Fichte's metaphysical commitments for their glaring epistemic insufficiency, this objection targets his philosophically more basic claim concerning the categorical force of reason's most basic requirement, suggesting that *this* claim is rationally unconvincing absent of any additional supporting evidence.

It is noteworthy that Fichte raises and replies to just this sort of objection in "Review of *Aenesidemus*," which contains the earliest published presentation of his system's first principles and the affiliated account of "what it means to say that reason is practical"⁶⁶—and which weaves all of that into a defense of Kant's moral theology, in which some broad and basic metaphysical commitments are justified as postulates of pure practical reason. Fichte phrases the objection thus:

Even the judgment *that* something is commanded is based on theoretical principles [that is, requires validation on epistemic grounds]. That which Kant first infers *from* the command has got to be already shown and decided *before* any command at all can be rationally accepted. It is far from being the case that the recognition of a command can provide the basis for the conviction that the conditions for its fulfillment actually do exist. On the contrary, that recognition can only follow upon this conviction.⁶⁷

Fichte notes that this objection "is assailing the actual foundation of Kantian moral theology, namely the primacy of practical over theoretical reason,"⁶⁸ and contends, somewhat cryptically, that criticisms of this kind reflect a "deficient grasp of the true difference between theoretical and practical philosophy."⁶⁹

The latter thought is rather more clearly conveyed when Fichte returns to this issue some years later in the *Divine Governance* essay: "A person who says, 'Before I can judge whether I ought to do something, I first have to know whether I can do it' either annuls the primacy of the ethical law itself. . . or else (if this is a purely speculative judgment) he totally

⁶⁶ EPW 75, SW 1:22.

⁶⁷ EPW 74–5, SW 1:21–2.

⁶⁸ EPW 75, SW 1:22.

⁶⁹ EPW 74, SW 1:21.

misunderstands the original progression of reason.”⁷⁰ The context makes it clear that for Fichte the “original progression of reason” is such that a hyperepistemic rational requirement (“the ethical law,” in the popularized passage just quoted) *is transcendentally prior to and therefore rationally outranks* all rational requirements of the narrowly epistemic sort: “this demand is rational and [is] the source and guideline of everything else which is rational.”⁷¹ More precisely: on Fichte’s account, it is “the supreme law of reason, i.e., the law of self-sufficiency . . . to which all the other laws of reason are subordinate and which provides the foundation for all of these other laws, at the same time that it determines them and limits them to the particular spheres in which they are valid.”⁷² We have already seen why and in what sense Fichte would say this. (1) On his account of the conditions for the possibility of experience, pure reason’s transcendentally most basic operation (the self-positing of the pure I) must articulate and assent to a highest, hyperepistemic ideal of unconditioned rational activity: untrammelled efficacy in the origination and instatement of pure ordering forms. And (2) this is an ideal to which both narrowly epistemic activity (e.g., positing a not-I as the explanatory ground of recalcitrant sensory data) and specifically ethical activity (in which reason’s basic commitment to reason’s unlimited efficacy guides the agent’s selection of more specific, subservient ends)⁷³ respond and approximate.

We may now be in a better position to mark some of the more radical consequences of these claims. First, in light of the foregoing, to require additional epistemic grounds indicative of the objective validity of pure reason’s highest ideal *does not qualify as a rational stance*. This is partly because this stance reverses “the original progression of reason” outlined above, inverting the actual order of priority among our rational ends and the associated rational requirements. In particular, in its suggestion that

⁷⁰ IWL 149, SW 5:184.

⁷¹ VM 101, SW 2:291.

⁷² IWL 39–40; SW 1:456–7.

⁷³ Concerning this side of Fichte’s position, see Michelle Kosch, “Agency and Self-Sufficiency in Fichte’s Ethics,” *Philosophy and Phenomenological Research* 91, no. 2 (2015): 348–80.

the rational acceptability of the ideal of unconditional rational activity finally depends on the strength of the supporting epistemic grounds, this position uncritically assumes that evidence-based justification is our highest rational priority, such that narrowly epistemic norms have unqualified rational authority. But for Fichte, as we have seen, the authority of any valid epistemic norm can only derive from its conforming and conducing to reason's ultimate, underived, hyperepistemic ideal (unconditioned rational activity), approximation to which is *the* supreme, superepistemic, rational requirement. To propose to overrule this requirement in service to an insistence on evidence-based justification is thus, in effect, to have misunderstood the basic nature and aim of rationality as such.

But there is another and, I think, transcendently more fundamental sense in which the aforementioned stance does not finally qualify as fully rational on Fichte's account. Note that the stance in effect suggests that we provisionally suspend, pending its objective validation on epistemic grounds, precisely the pure-rational commitment that is transcendently required (on Fichte's epistemology) in order to articulate and sustain *any possible framework for the truth-apt cognition of objectivity*—that is, any categorial framework the articulation of which allows for reference to a “not-I”: anything judged to be ontologically additional or extrinsic to the subject's own states and activities (regardless of whatever further descriptions it is also taken under). On Fichte's account, outside of this pure, categorially organized and normatively superintended frame of reference, there lies (for us) only an intellectually opaque and ethically indifferent mass of arational empirical data, the simple givenness of which does not suffice for any truth-apt judgment or reason-responsive volition. It seems, then, that to insist tough-mindedly that we set aside the upshot of Fichte's transcendental arguments and seek for more and better objective evidence of the validity of pure reason's ultimate ideal is, in effect, to imagine irrationally that we could conduct a rational investigation that is not confined to the cognitive standpoint constitutive of finite rationality as demarcated precisely by those transcendental arguments. Or, to view the same point from a different angle: given the transcendental role played, on Fichte's account, by the superordinate rational requirement to approximate to unconditioned rational

activity, it seems that there is and can be no *space of reasons* that is not finally superintended by that hyperepistemic requirement, no fully and properly rational frame of reference whose epistemic (and ethical) norms do not *derive from and subserve* that requirement. And in that case, it seems, there is no *rational* way around acknowledging that requirement's unqualified authority.

This is not to deny that the question "What evidence is there that the ultimate rational requirement is objectively and unconditionally prescriptive?" is in a certain sense intellectually unavoidable. Fichte openly acknowledges this: "argumentation possesses no immanent limit within itself. It freely proceeds into infinity, and it must be able to do so, for I am free in all of my expressions, and only I myself am able to set a limit for myself through willing"⁷⁴—that is, purposely to observe the proper order of priority among reason's highest, hyperepistemic end and its more specific and subservient epistemic aims. Perhaps the above question about evidence will still seem pressing, but on Fichte's epistemology that is to be expected. For one thing, to pose and pursue precisely this kind of question is at least *procedurally* rational, so to say: it is an instance of pure rational activity performed in service to an epistemic end. Moreover, given pure reason's inherent aspiration to unconditionality, we might feel as though we are called upon to raise questions and conduct critical trials without limit:

What can prevent speculation from asking such questions and from continuing to ask them indefinitely? What can I answer and where is there a point at which I could call a halt to its questions? I know, of course, and must concede this to speculation, that one can reflect on every determination of consciousness and produce a new consciousness of the first consciousness, that in doing this one always moves immediate consciousness one step higher and obscures and makes doubtful the first, and that this ladder has no highest rung.⁷⁵

⁷⁴ IWL 147, SW 5:182

⁷⁵ VM 70–1, SW 2:252–3.

As we saw above, according to Fichte's own metaphilosophy, that act whereby we bracket our given standpoint, suspend its constitutive commitments, and initiate higher-order critical reflection upon that entire outlook is methodologically definitive of and foundational for philosophy properly so called. Interestingly, however, he denies that we are philosophically entitled to utilize this device without restriction. Note that if we were so entitled, then, evidently, to realize philosophy's proper possibilities would *not* be to articulate rationally a life-clarifying, action-guiding outlook and orientation; instead, it would be "rationally" to revoke or relativize any such outlook or orientation lest we "uncritically" acquiesce in it. That is to say, if the task proper to philosophy is to slip the leash of our everyday outlook and its incipient "dogmatism" *just in order* to embark upon an in-principle interminable series of further bracketings and suspensions, then, insofar as we philosophize competently, we "plunge into the realm of what is unbounded and ungrounded, and . . . dispense absolutely with any firm standpoint whatsoever . . . We remain upon a boundless ocean where every wave is propelled forward by yet another."⁷⁶

Obviously Fichte roundly rejects this conception of the philosophical project—and indeed, I suggest, of *the project proper to finite rationality as such*, as disclosed via his transcendental epistemology. Note that his rejection of the above approach is based not so much on its neoskeptical upshot as on its substantively irrational *inception*. As indicated above, such an enterprise seems to qualify as formally rational in its procedure. By Fichtean lights, however, it fails to be substantively rational in its foundation or inception, which principally consists in a tacit refusal (i) to respect the proper order of priority among reason's ownmost goals, and—what for Fichte is the same thing, only stated in different terms—(ii) to remain within the standpoint and space of reasons proper to finite rationality.

For Fichte, to remain in that place is to respect the rightful order of rational norms, which means to acknowledge the real and unconditional prescriptivity of the ideal of unconditioned rational activity. "Here lies

⁷⁶IWL 148, SW 5:182. Cf. n.3, above, on the related issue of ontotheology in Fichte.

that which sets a limit to the otherwise unbridled flight of argumentation,” he says, “that which binds the mind because it binds the heart.”⁷⁷ Acknowledging the unconditionally binding authority of reason’s ultimate, hyperepistemic end “binds the heart,” and in turn this acknowledgment “binds the mind” by rationally, albeit nonepistemically, enjoining assent to those propositions that affirm the basic metaphysical conditions for that ideal’s unqualified validity. The nonepistemic grounds for the indicated assents are (i) rationally unassailable, given that the issues involved are epistemically intractable, and (ii) rationally irrevocable, insofar as pure reason is practical. “There is no firm standpoint except the one just indicated,” Fichte therefore says, “and it is based not upon logic, but upon one’s moral disposition or sentiment.”⁷⁸ The affiliated ethico-metaphysical outlook, despite being a fit object of firm assent, is not the upshot of “any sort of logically coercive thought.”⁷⁹ Instead it is, in Fichte’s words, “the way of thinking which I have carefully, intentionally, and with consideration chosen from among other possible ways of thinking, because I recognized it to be the only one appropriate to my dignity and vocation.”⁸⁰

Given Fichte’s epistemology *plus evidentialism*, the indicated outlook would prove rationally unacceptable. But on Fichte’s account, precisely that outlook uniquely admits of “deduction”—in the Kantian sense connoting strict legitimation—“*from the very nature of reason.*”⁸¹ It is therefore not just rationally acceptable, but even, in the philosophically final analysis, rationally unavoidable. As we saw above, for Fichte (and strictly from the higher-order transcendental standpoint proper to philosophy) there is no rational way around acknowledging the unqualified authority of reason’s intrinsic commitment to unconditioned rational activity. Consequently (given the epistemic intractability of the issues involved, and per the reasoning outlined above), there is,

⁷⁷ IWL 147, SW 5:182.

⁷⁸ IWL 148, SW 5:182. Cf. VM 94, SW 2:283.

⁷⁹ IWL 147, SW 5:182.

⁸⁰ VM 74, SW 2:256.

⁸¹ IWL 166, SW 5:386.

for the philosopher, no rational alternative to assenting to the indicated metaphysics—for “only on this view can we fulfill our vocation.”⁸² The hyperepistemic rational requirement that nonepistemically justifies the indicated beliefs determines and delimits not only *when* we may rationally believe on insufficient evidence (namely, only where the beliefs in question concern the metaphysical preconditions of the unqualified validity of reason’s highest requirement) but also *what* we may then rationally believe (namely, only those propositions that affirm that these preconditions obtain). “This belief,” Fichte says,

should not be represented as, so to speak, an arbitrary assumption one may adopt or not adopt as one pleases, that is, as a free decision to consider true whatever the heart wishes and to do so because this is what it wishes. Nor should this belief be represented as a hope that supplements or takes the place of sufficient (or insufficient?) grounds of conviction. What is grounded in reason is purely and simply necessary, and what is not necessary is—precisely for this reason—contrary to reason.⁸³

Clearly, then, such “belief” or “faith” (*Glaube*) does not reflect an irrational fideism, nor does it allow any arational factors to do any justificatory work. In light of the parallels with Kant’s moral theology, and given the phrasings Fichte favors in his popular presentations, it is tempting to label his view as a form of *practical* nonevidentialism, according to which moral requirements grounded in reason can justify beliefs for which the evidence is insufficient. But while that description would not be entirely inaccurate, it might somewhat oversimplify Fichte’s position, insofar as it does not exactly track the exceptionally abstract—that is, not exclusively or ultimately “moral”—way in which he actually understands the operative rational requirement. As Fichte sees it, reason’s highest end is not just hyperepistemic but *also hyper-ethical*: “the complete annihilation of the individual and the fusion of the latter into the absolutely pure form of reason or into God is indeed

⁸² VM 71, SW 2:253.

⁸³ IWL 144, SW 5:179.

the ultimate goal of finite reason.”⁸⁴ Thus, the rational requirement that underwrites his nonevidentialism, while definitely hyperepistemic, is not narrowly moral or ethical in its origin or its ultimate import. For that reason, I suggest, we do best to think of Fichte’s position as a *pure-rational nonevidentialism*, understanding “pure-rational” in the very rarefied sense that derives from the first principles of his transcendental philosophy.⁸⁵

Steven Hoeltzel is Professor of Philosophy at James Madison University. A specialist in Kant and post-Kantian idealism, his recent publications include “Transcendental Idealism and Theistic Commitment in Fichte,” in *The Palgrave Handbook of German Idealism* (Palgrave Macmillan, 2014), and “Non-Epistemic Justification and Practical Postulation in Fichte,” in *Fichte and Transcendental Philosophy* (Palgrave Macmillan, 2014). He is also the editor (with H. Kim) of *Kant, Fichte, and the Legacy of Transcendental Idealism* (Lexington Books, 2014). His recent writings focus on the philosophical ramifications of Fichte’s nonstandard metaphilosophy and post-Kantian conception of reason.

⁸⁴ SE 143, SW 4:151.

⁸⁵ An early draft of this chapter was presented at the 2015 Conference on Transcendental Philosophy and Metaphysics in Osaka, Japan. This revised version is much indebted to subsequent conversations with Robert Stern and correspondence with Benjamin D. Crowe.

5

Transcendental Philosophy as “Therapy of the Mind”

Fichte’s “Facts of Consciousness” Lectures

Benjamin D. Crowe

The last several decades have seen a flourishing of international scholarly interest in Fichte’s thought, considered both in its own right and as a decisive moment in the development of transcendental philosophy. Yet, at least in anglophone circles, where the revival of interest in his has been most striking, Fichte’s work after his departure from Jena in 1799 largely remains terra incognita. In the interests of correcting this situation, my goal in this chapter is to examine the way in which Fichte takes up transcendental philosophy in this later period.¹ While the significance and extent of Fichte’s later revisions of the position he developed during

¹ Fichte’s commitment to “critical” or “transcendental” idealism in the early, more well-known, period of his career is, I take it, not a point of controversy. See, for example, the following remark in the First Introduction to the 1797 *Attempt at a New Presentation of the Wissenschaftslehre*: “The *Wissenschaftslehre* wishes to establish a complete transcendental idealism” (IWL 34, GA 1/4:208).

B.D. Crowe (✉)

Department of Philosophy, University of Utah, Salt Lake City, USA
and Boston University, Boston, USA

the Jena years are matters of much scholarly debate, it is apparent that he takes himself to be a transcendental idealist even in the final period of his career. For instance, one of many comments to this effect can be gleaned from his lectures on logic from the summer of 1812: “all knowing has itself as an object, and nothing else at all. . . . This is the content of the WL [*Wissenschaftslehre*], of transcendental idealism.”²

Throughout his career, Fichte belonged firmly in the camp of those of Kant’s successors who, in Paul Franks’s words, “think that transcendental philosophy is intelligible only to someone who occupies an appropriate *standpoint*, distinct and in some sense opposed to the empirical standpoint pertinent to ordinary life, natural science, and non-transcendental philosophy.”³ The importance of occupying the right standpoint from which to derive the system of transcendental philosophy is the centerpiece of Fichte’s lectures on the “Facts of Consciousness” (delivered between October 1810 and January 1811), one of the most complete lecture series that he presented in the last part of his career.⁴ The lectures progress through a series of what might be termed “quasi-transcendental arguments” (the course of which I will describe at the conclusion of this chapter) that involve a move from a particular structure (or “fact”) to a further one that embeds it. These moves are not, as in a transcendental argument standardly conceived, aimed at answering the *quaestio juris* about pure a priori concepts; rather, the goal is to establish the standpoint from which the *quaestio juris* can be adequately taken up. The lectures also have a definite target: Fichte aims to undermine not only the skeptical misconstruals

² GA 2/14:48.

³ Paul Franks, “Serpentine Naturalism and Protean Nihilism: Transcendental Philosophy in Anthropological Neo-Kantianism, German Idealism, and Neo-Kantianism,” in *The Oxford Handbook of Continental Philosophy*, ed. Brian Leiter and Michael Rosen (Oxford: Oxford University Press, 2007), 251.

⁴ Fichte apparently delivered four courses in Berlin under this title. The other three were held in the summer term of 1811 (between April 22 and July 12), the winter term of 1811–12 (between October 21 and December 20), and the winter term of 1812–13 (between January 4 and February 4). We possess extant texts of the winter 1810–11 course (in Fichte’s MS), the winter 1811–12 course (in various student transcripts), and the winter 1812–13 course (in various student transcripts).

of the *Wissenschaftslehre* as hyperbolic subjectivism or nihilism, but also the dogmatism resurgent in the early post-Kantian period.

My discussion proceeds as follows. First, I explore how, despite its relatively late date in Fichte’s career, this lecture course can be helpfully framed by reference to Kant’s characterization of his project in the Transcendental Deduction. Next, I describe how Fichte remained dissatisfied with Kant on account of the latter’s vestigial empiricism and naturalism, which, according to Fichte, paved the way for a fundamental misconstrual of his own system of transcendental philosophy. Finally, having shown how Fichte locates the incompleteness of Kant’s project in a tendency within the “natural” or “vulgar” consciousness, I describe how the “Facts of Consciousness” lectures involve the appropriation of a therapeutic conception of philosophy, which is realized in the lectures’ “quasi-transcendental” argumentation. The *terminus ad quem* of this philosophical therapy is the free recognition that, to use Fichte’s words, “there is a knowing [*Wissen*] that is actually independent [*selbständig*] and exists in fact; this knowing is a free and independent life.”⁵

Kantian Background

Beginning with his “conversion” to critical philosophy around 1790, Fichte consistently cites Kant as the progenitor of transcendental philosophy, and thus of the basic philosophical orientation of his own thought.⁶ It should not be a surprise, therefore, that the “Facts of Consciousness” lectures can be partly illuminated by

⁵ StA 1:393.

⁶ As with Fichte’s later intellectual development, some interesting questions can be raised about the relationship between his youthful intellectual outlook and that which eventuated from his encounter with Kant’s writings. Rainer Preul, for example, makes the case that Fichte’s earliest thought provided fertile ground for the reception of Kant. See *Reflexion und Gefühl: Die Theologie Fichtes in seiner vorkantischen Zeit* (Berlin: de Gruyter, 1969). For an exploration of how tensions in Fichte’s earliest theological thinking get taken up during his Jena period work, see my “Fichte on Faith and Autonomy,” *British Journal for the History of Philosophy* 21, no. 4 (2013): 733–53.

having recourse to a Kantian precedent. In this case, some of Fichte's comments about the project of the lectures furnish useful clues. At the beginning of the lectures, he observes that the "essence of all science" consists in the elevation, "through thought, from something that is sensibly perceived to its supersensible ground."⁷ In philosophy, this trajectory begins with "the perception of knowing [*Wissen*] through the inner sense." "In these lectures," writes Fichte, "we are concerned with the first piece of this science, with the phenomena, which we want to systematically observe."⁸ Elsewhere, he draws a distinction between what is being undertaken in these lectures and philosophy (or *Wissenschaftslehre*) proper, which is the "demonstration [*Nachweisung*] of the necessity of the forms" of knowing.⁹ August Twesten's transcript of the introductory lectures delivered immediately prior to the lecture course under consideration here likewise labels their topic as "phenomena," with the following interesting proviso: "what these lectures will contain is in no way science or philosophy, but rather experience [*Empirie*], which, however, is distinct from common experience in that it describes not something that is immediately given in perception, but rather, as will be shown, something that is produced by freedom."¹⁰

The distinction that Fichte is drawing here is between an investigation that *justifies* or *demonstrates the necessity of* the "forms of knowing" and one that approximates an empirical or descriptive presentation of these forms. In framing the lectures in this way,

⁷ StA 1:229.

⁸ StA 1:229.

⁹ StA 1:390. In his lectures on the *Wissenschaftslehre* from 1810, Fichte argues that philosophy is *deductive* and *genetic*, and that only as such can it demonstrate the necessity of something (StA 1:57–8). Similarly, in the 1812 lectures on the *Wissenschaftslehre*, he urges that it is "supra-factual [*überfaktisches*]," that is, that it is not about any actual configuration of appearance, but rather about how appearance *must* be constituted (GA 2/13:99). The *Wissenschaftslehre*, he later points out again, "deals with absolute forms, irrespective of what these might signify in actual consciousness or with how they might happen to appear there" (GA 2/13:133).

¹⁰ StA 1:220.

Fichte’s comments echo a well-known passage in the *Critique of Pure Reason*, where Kant contrasts his own project with that of the “famous Locke”:

Such a tracing of the first endeavors of our power of cognition to ascend from individual perceptions to general concepts is without doubt of great utility, and the famous Locke is to be thanked for having first opened the way for this. Yet a *deduction* of the pure a priori concepts can never be achieved in this way; it does not lie down this path at all, for in regard to their future use, which should be entirely independent of experience, an entirely different birth certificate than that of an ancestry from experiences must be produced. I will therefore call this attempted physiological derivation, which cannot properly be called a deduction at all because it concerns a *quaestio facti*, the explanation of the *possession* of a pure cognition. It is therefore clear that only a transcendental and never an empirical deduction of them can be given, and that in regard to pure a priori concepts empirical deductions are nothing but idle attempts, which can occupy only those who have not grasped the entirely distinctive nature of these cognitions.¹¹

Kant then revisits the distinction drawn here in §§26–7 of the B edition, first observing that “categories are concepts that prescribe laws a priori to appearances,” and so “are not derived from nature and do not follow it as their pattern.”¹² The *necessity* of the agreement between these purely a priori elements of cognition with experience can only be secured, Kant thinks, through “a system of the *epigenesis* of pure reason.” The key feature of this system is the claim that “the categories contain the grounds of the possibility of all experience in general from the side of the understanding.”¹³ As Kant sees it, any alternative concedes too much to the skeptic, that is, that the categories are “subjective” and so “lack the

¹¹ B118/A86–B119/A87.

¹² B163.

¹³ B167.

necessity that is essential to their concept.”¹⁴ No doubt with Hume in mind, Kant illustrates the point in reference to the concept of causation:

For, e.g., the concept of cause, which asserts the necessity of a consequent under a presupposed condition, would be false if it rested only on a subjective necessity, arbitrarily implanted in us... I would not be able to say that the effect is combined with the cause in the object (i.e., necessarily), but only that I am so constituted that I cannot think of this representation otherwise than as so connected; which is precisely what the skeptic wishes most, for then all of our insight through the supposed objective validity of our judgments is nothing but sheer illusion, and there would be no shortage of people who would not concede this subjective necessity (which must be felt) as their own; at least one would not be able to quarrel with anyone about that which merely depends on the way in which his subject is organized.¹⁵

Fichte takes this warning very much to heart, particularly in light of the fact that his own view was very often taken to be subjectivist in an absurdly exaggerated sense. As I will describe below, it is precisely the failure to adopt something like an “epigenetic” outlook on reason (or “knowing”) that leads to this misunderstanding. Fichte is committed to Kant’s assertion that only a purely a priori “birth certificate” serves to give rational grounding to our fundamental concepts. In a letter of June 23, 1804, in response to a request for a “concept of transcendental philosophy,” Fichte argues that a philosophy that is “strictly scientific” proceeds not through “empirical self-observation” but “purely and simply a priori, without the assistance of empirical perception.”¹⁶

Fichte’s distinction between philosophy proper, which demonstrates the *necessity* of the forms of knowing, and the presentation of the *phenomena* of knowing, maps readily onto Kant’s distinction between a “transcendental” and “physiological” (or empirical) derivation of the

¹⁴ B167–8.

¹⁵ B168.

¹⁶ GA 3/5:246–7.

categories. (Recall that Fichte qualifies the “Facts of Consciousness” as “experience” [*Empirie*].) Kant’s characterization of the Lockean model as tracing out the *development* of our power of cognition is similarly paralleled by the actual course of Fichte’s lectures. Not only do the lectures themselves read like a developmental narrative, but they also contain broadly empirical descriptions of human cognitive development.¹⁷ Indeed, Fichte maintained a lifelong fascination with what we would now term developmental psychology; in particular, he enthusiastically received the empirical approach to childhood education developed by J. H. Pestalozzi (1746–1827).¹⁸ Fichte’s remarks at the conclusion of the 1810–11 lectures make their biological-developmental character explicit: “a presentation [*Darlegung*] of the facts of consciousness would thus be as it were a *natural history of the development of this life* [of knowing].”¹⁹

¹⁷ Fichte goes into some detail describing the distinction between the stage of cognitive development of a child and that of an adult during his discussion of the faculty of imagination. He contrasts, for example, the fairly limited capacity of a young child for “abstraction from sensible impressions” with that of Archimedes, who was allegedly not distracted from his calculations by a massive assault on his home city by the Roman army (StA 1:244). He goes on to develop the contrast further, comparing the way a child perceives a particular plant with the way a naturalist (*Naturforscher*) or botanist possesses a kind of “free perception under the guidance of a purposive concept [*Zweckbegriff*].” The latter “can undertake the observation in accordance with a certain order, can linger over certain parts as long as he likes until he is aware that he has seen [the plant] correctly” (StA 1:247).

¹⁸ An early reference to Pestalozzi’s *Leonard und Gertrude* (1781–83) can be found in a fragmentary text from 1788: GA 2/1:103–10. In a letter to his wife, written from Königsburg in June 1807, Fichte requests that she forward him Pestalozzi’s more recent work, *Wie Gertrud ihre Kinder Lehrt* (1801), describing the latter’s system as being not only “the true means of healing sick humanity” but also as “the only means for making [humanity] fit for understanding the *Wissenschaftslehre*” (GA 3/6:121).

¹⁹ StA 1:394, emphasis added. The appeal to natural history as a model for philosophical method can also be found in Friedrich Schlegel’s review of Niethammer’s *Philosophisches Journal*, first published in 1797. Interestingly, Schlegel contrasts the method of natural history with the *deductive* method characteristic of transcendental philosophy. Schlegel was living in Jena during the time that he composed this review, and interacted closely with Fichte throughout his time there. Whether or not Fichte had Schlegel’s appeal to natural history in mind here in the “Facts of Consciousness” is unclear. For a discussion of Schlegel’s review, see Dalia Nassar, *The Romantic Absolute: Being and Knowing in Early German Romantic Philosophy, 1795–1804* (Chicago: Chicago University Press, 2014), 96–7.

The concept of epigenesis to which Kant appeals in the passage from the B Deduction derives from developmental biology (more specifically, embryology) as well. In essence, an epigenetic model of biological development involves inherent capacities that unfold through a spontaneous development; that is, given the inborn capacities, organisms develop themselves.²⁰ Fichte was familiar with some of the leading work on epigenesis by J. F. Blumenbach (1752–1840), whose concept of a *Bildungstrieb* appears, for example, in the 1798 *System of Ethics*.²¹

At the same time, it is clear that the “Facts of Consciousness” lectures are not to be taken as an actual “natural history.” For one thing, the comments reported by Twesten contrast the experience of something *given* in perception with that of something *produced* by freedom. For another, there are Fichte’s remarks at the beginning of the lectures, namely, that they involve “an artificially conducted experiment” that presents knowing “not in its immediate, living being, but rather only in the image of this being.”²² Fichte’s procedure is not purely descriptive. Rather, as he explains, it involves a kind of purpose-driven selection or “systematic observation” that often “requires that we adopt a particular, contrived preconception [*künstlichen Vorsicht*] so that consciousness responds directly to the questions that we put to it.”²³ Thus, if the “Facts of Consciousness” do not present a strict transcendental

²⁰ There is a large and growing literature on the significance of this passage in the B Deduction as well as other, similar discussions throughout Kant’s corpus. Günter Zöllner brings out the importance of the earlier, rationalist context in “From Innate to *A Priori*: Kant’s Radical Transformation of a Cartesian–Leibnizian Legacy,” *Monist* 72, no. 2 (1989): 222–35. Another important treatment is found in Hans Ingensiep, “Die biologischen Analogien und die erkenntnistheoretischen Alternativen in Kants *Kritik der reinen Vernunft*,” *Kant-Studien* 85, no. 4 (1994): 381–93. Somewhat more recently, Phillip Sloan provides a more comprehensive contextualization and interpretation in “Performing the Categories: Eighteenth-Century Generation Theory and the Biological Roots of Kant’s *A Priori*,” *Journal of the History of Philosophy* 40, no. 2 (2002): 229–53. By far the most complete and detailed discussion, however, is found in Jennifer Mensch, *Kant’s Organicism: Epigenesis and the Development of Critical Philosophy* (Chicago: Chicago University Press, 2013). Ch. 7 focuses on the idea of reason as “self-born” or “self-generated.”

²¹ Fichte may have been aware of Blumenbach’s ideas through the mediation of Kant, who discusses Blumenbach sympathetically in §§80–1 of the *Critique of the Power of Judgment*.

²² StA 1:229.

²³ StA 1:229.

deduction (to revert for a moment to Kant), neither do they present a genuinely “physiological derivation.” At the same time, it is nevertheless clear that Fichte regards the presentation of the “facts of consciousness” as an indispensable element of transcendental philosophy. As I will explain in the sequel, the positive aim of the lectures is to conduct the audience to the free acknowledgment of the distinctive point of view that is required for the successful execution of transcendental argumentation in the *Wissenschaftslehre* itself.

“A New Human Being Must be Born”

Somewhat confusingly, Fichte only articulates the ultimate purpose of the lectures, and their central role in his transcendental project, at their conclusion. The basic claim of transcendental idealism (as Fichte understands it) is put this way: “there is a knowing [*Wissen*] that is actually independent [*selbständig*] and exists in fact; this knowing is a free and independent life.”²⁴ It is hard to overlook the parallel between the way Fichte puts this key point and Kant’s appeal to biological epigenesis as a model for the transcendental deduction of the categories. Also like Kant, Fichte wants to avoid any hint that the forms of knowing are simply features of the subjective constitution of an individual. This “free and independent life” is not a mode or property of a particular mental substance, “substantial person,” or “substrate” (*Träger*).²⁵ Somewhat notoriously, it is less apparent how this “free and independent life” can be construed positively. Be that as it may, the important point here is that, for Fichte, it is not enough simply to understand this assertion; rather, “one must also actually have this thought, and must have developed his faculty of thought to the point that he can have this thought, if only to then subject it to examination”²⁶ That is, philosophy proper must proceed *from* this standpoint, rather than having it as its conclusion.

²⁴ StA 1:393.

²⁵ StA 1:393.

²⁶ StA 1:393.

But, as Fichte is well aware, this standpoint is something that has to be adopted. The free and independent life of knowing “begins in a certain boundedness [*Gebundenheit*] of its freedom.”²⁷ Thus, “its progress and the course of its life consists in it liberating itself from this boundedness, whereby it may well fall prey to a different, lesser boundedness, from which it must once again liberate itself, etc. In short, the course of its life is a progressive raising of its life to a higher freedom.”²⁸ It turns out that the “natural history” that is supposedly traced in the “Facts of Consciousness” lectures just is the law-governed “progress and course” of this liberation. It moves upwards from the lowest point, the point at which life is simply given without previous development: “this point, this *terminus a quo* of the history is external perception.”²⁹ Recall, however, that what is being traced in the lectures is not the *actual* natural history of reason. Instead, it is the reason of Fichte’s *audience* that is being led, through a nested series of intuitions, to the perspective of the “free and independent life of knowing.” It is as though the lectures recreate or, better, re-enact the developmental process of reason *überhaupt*. As I will argue in the next section, this re-enactment is best thought of as a kind of philosophical therapy.

Again, Fichte readily acknowledges that the basic standpoint from which transcendental philosophy proceeds is not something that can be taken for granted, nor is it necessarily all that easy to adopt. This observation is most salient in the context of the debates that broke out in the 1790s regarding the genuine upshot of Kant’s transcendental turn. In the lectures under consideration here, Fichte writes that “the philosophers that have been opposed to us in this regard have not allowed consciousness to be valid as an appearance that stands on its own feet . . . Kant is the pioneer and first person to discover our manner of treating it.”³⁰ In other words, what has happened is that philosophers have failed to grasp the genuine lesson of Kant, which,

²⁷ StA 1:393.

²⁸ StA 1:393–4.

²⁹ StA 1:394.

³⁰ StA 1:321.

for Fichte, is articulated by the recognition that “there is a knowing [*Wissen*] that is actually independent [*selbständig*] and exists in fact; this knowing is a free and independent life.”³¹ Fichte’s comments make it clear that his immediate target is Schelling, whose 1806 polemical piece, *Exposition of the True Relation of Naturphilosophie to the Improved Fichteian Doctrine* (*Darlegung des wahren Verhältniß der Naturphilosophie zu der verbesserten Fichte’schen Lehre*), is referenced throughout these lectures. The title of the lectures, however, indicates a broader target as well. The reception of Kant in the 1790s and beyond involved not just the rise of German idealism, but also of what has been called “anthropological Neo-Kantianism.” Fichte’s writings from the 1790s are peppered with critical references to such “so-called Kantians,” whose appeal to the justificatory force of “facts of consciousness” (*Thatsachen des Bewußtseyns*) is contrasted with a genuinely *transcendental* approach to philosophy. Fichte’s opponents here included C. C. E. Schmid, Jakob Friedrich Fries, Gottlob Ernst Schulze, and Johann Friedrich Herbart.³²

³¹ StA 1:393.

³² Fichte refers to the “so-called Kantians” and their appeals to the “facts of consciousness” throughout the 1797 *Attempt at a New Presentation of the Wissenschaftslehre*. See, for example, IWL 9, GA 1/4:187; IWL 69–70, GA 1/4:237–8. Fichte makes several remarks in a long footnote to the “First Introduction” (IWL 15, GA 1/4:192) that resonate with some of his comments in the “Facts of Consciousness” lectures:

This form of “Kantianism” is the most fantastic monster that human fantasy has ever engendered, and it does little credit to the perspicuity of its defenders that they fail to realize this. Furthermore, it can be easily proven that the only thing that recommends this philosophy is that it allows people to dispense with all serious speculation and allows them to believe that they have been granted a royal patent authorizing them to continue cultivating their beloved and superficial empiricism.

Fichte’s polemic against the “so-called Kantians” reached its unedifying nadir with the famous “act of annihilation” aimed at C. C. E. Schmid (see EPW 316–35, GA 1/3:235–47). Schmid’s appeals to “facts of consciousness” figure prominently in Fichte’s critique. For more recent examinations of this distinctive line of post-Kantian thought, see Terry Pinkard, *German Philosophy 1760–1860: The Legacy of Idealism* (Cambridge: Cambridge University Press, 2002), especially ch. 8; see also Paul Franks, “Ancient Skepticism, Modern Naturalism, and Nihilism in Hegel’s Early Jena Writings,” in *The Cambridge Companion to Hegel and Nineteenth-Century Philosophy*, ed. Frederick C. Beiser (Cambridge: Cambridge University Press, 2008), 52–73.

Fichte grants to these opponents that Kant himself occasionally departs from the transcendental, deductive method, and that there are traces of vestigial empiricism even in the work of the originator of transcendental philosophy. In the present lectures, Fichte argues that Kant failed to purge sufficiently his own thought from the specter of empirical or subjective idealism, for the simple reason that the latter appeals to “common sense,” to *facts* supposedly available to inner intuition, in order to establish claims about the “universal rational form of the I.”³³ As Fichte puts it in the 1804 letter quoted previously, Kant fell prey to the same “wisdom in which we are all brought up [*aufgewachsen*],” namely, “*that experience, observation, the empirical remain what is highest and ultimate, and that we can never transcend them.*”³⁴ The outcome of this “wisdom” is precisely the confusion that Kant hoped to avoid in the B Deduction, namely, the notion that (what Fichte calls) the “forms of knowing” are features of the constitution of a subject, or empirically given “facts of consciousness.” The danger here is the conflation of “reason” or “knowing” as such with “my reason” or “my knowing.” Fichte envisions various equally bad outcomes from this conflation. One, embodied for him by the work of the “anthropological Neo-Kantians,” would be the naïve notion that one could appeal to a “fact of consciousness” to ground somehow or justify the a priori forms of knowing.³⁵ Another consequence of this conflation is demonstrated by readings of the *Wissenschaftslehre* as being perniciously subjective. The most famous of these originated from Jacobi, who observes in his 1799 attack on Fichte that “every science . . . is an object-subject according to

³³ StA 1:324.

³⁴ GA 3/5:245.

³⁵ Fichte had long worried about this direction in post-Kantian thought. Some of his earliest writings evince this concern by appealing to a “fact of consciousness.” See, for example, his 1792 review of Friedrich Heinrich Gebhard’s *Ueber die sittliche Güte aus uninteressirtem Wohlwollen* (GA 1/2:15–30), or, more famously, the discussion of Reinhold in the *Aenesidemus* review of the same year (EPW 59–77, GA 1/2:41–67). In the 1798 *System der Sittenlehre*, Fichte similarly argues that a properly philosophical account of moral consciousness cannot rest merely on the *fact* of the sense of moral obligation (GA 1/5:110). For a discussion of Fichte’s suspicions in this regard, see Frederick Neuhouser, *Fichte’s Theory of Subjectivity* (Cambridge: Cambridge University Press, 1990), 25–9.

the prototype of the I,”³⁶ and hence that “[Fichte’s] own pure reason, everywhere present in him, does not prevent him from thinking and assuming the most absurd things.”³⁷ On Fichte’s view, the effort to avoid subjectivism is what led Schelling down the slippery slope to *Naturphilosophie*, which he regards as nothing but a resurgent dogmatism.

At the end of the lectures, commenting on the erroneous claim that knowing requires a substrate, Fichte again refers to Kant’s supposed limitations, as well as those of the “anthropological Neo-Kantians”:

Indeed, Kant did not frankly express this proposition with such bald language [*dürren Worten*] as we have; but in fact he said nothing without assuming it, and his writings remain a mess of contradictions. The philosophical public did not make this assumption, and so actually found [either] nothing, or a mess of contradictions, in these writings. (It is admittedly a puzzle how it could be that some whose faculty of thought is in this condition have nevertheless found wisdom in this doctrine and are even moved to expound and propagate it.)³⁸

The assumption or assumptions that underwrite this supposed “mess of contradictions” turn out to be quite natural; indeed, Fichte argues earlier that this situation simply is that of the “natural person” who stands at the *terminus a quo* of the “natural history” traced by the lectures:

In this way it also becomes comprehensible that someone who is told that knowing is absolutely [*schlechtweg*] a self-sufficient life knows of no other way to understand this than that self-consciousness is what is being talked about. According to the necessary laws of thought he is unable to transcend individuality to think of life in its unity. From the concentration of life to

³⁶ F. H. Jacobi, *The Main Philosophical Writings and the Novel Allwill*, trans. and ed. George di Giovanni (Montreal: McGill-Queen’s University Press, 2009), 505.

³⁷ Jacobi, *The Main Philosophical Writings*, 514.

³⁸ StA 1:395.

one point, which is itself an absolute fact [*Faktum*], everything is factual. The natural human being is a merely historical intellect [*historisch Intelligenz*] who indeed understands facts and can copy them in the reproductive imagination, can posit one in place of another and transpose them with one another. But with this he has reached the absolute bottom and the limits of his horizon. But when it comes down to no longer substituting facts for facts, but rather elevating oneself absolutely beyond all facticity according to its absolute form to the absolute ground of the same via pure thinking, there the capacities of the natural human being are exhausted[;] the natural human being must die and the new human being be born.³⁹

Throughout writings from this period in his career, Fichte identifies what he calls here the “historical intellect” as the primary barrier to the positive reception of the *Wissenschaftslehre*.⁴⁰ Echoing the quasi-religious language of the passage quoted immediately above, Fichte comments in the 1811 lectures on the *Wissenschaftslehre* that “what matters is to distinguish what natural consciousness does not distinguish, what a human being by nature has a tendency not to distinguish. Hence, here there is a counter-striving [*widerstreben*], a tearing oneself free and, in fact, a new creature.”⁴¹ Or again, in the lectures on the *Wissenschaftslehre* from 1812, Fichte comments on the “difficulty” of genuinely transcendental philosophy in similar terms: “its task is to raise to consciousness what *must* remain completely invisible within vulgar consciousness: an expansion of the realm of light, a seeing that is contrary to nature.”⁴²

This consideration of Fichte’s criticisms of the “natural consciousness” or the “historical intellect” illuminates the position that the lectures on the “Facts of Consciousness” occupy within the project of transcendental philosophy. As described above, the course of the lectures, as well as

³⁹ StA 1:348–9.

⁴⁰ To take one example from the 1811 lectures on the *Wissenschaftslehre*, see Fichte’s remarks on the view that only the historical or factual can actually be exhibited, and that any philosophical claim must be factually demonstrated: StA 2:18.

⁴¹ StA 2:21.

⁴² GA 2/13:49.

Fichte’s comments concerning their goal and method, call to mind Kant’s discussion of the “physiological” approach to the pure concepts of the understanding. At the same time, Fichte disavows any straightforwardly empirical or naturalistic interpretation of his enterprise. Moreover, it is precisely the vestiges of Lockean “physiology” that, for Fichte anyway, have compromised the integrity of the Kantian project. By titling this lecture course as he does, Fichte signals his intention of taking on the “anthropological Neo-Kantians” and other assorted “dogmatists,” and in so doing bringing to fruition the genuine philosophical promise of Kant’s “Copernican revolution.”

The Therapy of the Mind

Fichte’s insistence on the necessity of adopting the proper standpoint for transcendental philosophy and his critical attitude towards the “historical intellect” converge in the “Facts of Consciousness” lectures. The requisite standpoint is, as I have argued above, a kind of radicalization of the epigenetic model that Kant urges in the B Deduction. The radical nature of Fichte’s model drew the censure not only of the “anthropological Neo-Kantians,” but, famously, of Kant himself (in 1799). At the same time, Fichte’s concerns and aims naturally led him to appropriate a venerable (indeed, ancient) way of thinking about the sort of philosophical enterprise at issue in the “Facts of Consciousness” lectures. The “quasi-transcendental” arguments of these lectures, aimed at establishing the point of view from which the deduction of the necessity of the forms of knowing must proceed, are exercises in *philosophical therapy*.

Fichte’s engagement with the idea of philosophy as therapy emerges in another series of lectures, given shortly prior to the foundation of the new University of Berlin, concerned with an “introduction to philosophy as a whole.” After commenting that the essence of philosophy is “reflection [*Besonnenheit*],” Fichte briefly takes stock of his own philosophical trajectory and the still unfinished quality of the *Wissenschaftslehre*. “I aspire only to educate [*erziehen*] to a certain point of view [*Ansichten*], to sharpen the critical and ideal sense as such, in

order to overcome blind, rigid faith [*blinde Stokgläubigkeit*] in the empirical. Thus [I aspire to] erect a genuine *medicina mentis*.”⁴³

The concept of philosophy as therapeutic has an ancient pedigree; among the ancient sources of the idea in Fichte’s case, Cicero is the most likely candidate. In *Tusculan Disputations* III.3.6, Cicero calls philosophy “an art of healing the soul [*animi medicina*].” Arthur Schopenhauer’s transcript of Fichte’s 1811 lectures “Über das Studium der Philosophie” indicate that Fichte cites Cicero directly, most particularly the defense of philosophy in the *Tusculan Disputations*. This Ciceronian idea had been enthusiastically taken up once again in the early modern period. For example, both Francis Bacon and his successors in the Royal Society argued that natural philosophy is best viewed as a cure for intellectual diseases like superstition. Regarding Bacon specifically, Peter Harrison observes how “the medicine of the mind is offered [in both *The Advancement of Learning* and the *Novum Organum*] as the practical component of a moral philosophical project.”⁴⁴ Bacon’s famous discussion of the “idols of the mind” quite clearly instantiates the therapeutic element of his project.⁴⁵ Fichte’s stated desire to overturn what he calls *Stokgläubigkeit* (*sic*) evokes this tradition.

⁴³ StA 1:4.

⁴⁴ Peter Harrison, “Francis Bacon, Natural Philosophy and the Cultivation of the Mind,” *Perspectives on Science* 20, no. 2 (2012): 139–58, 147.

⁴⁵ Corneanu and Vermeir make the following relevant observation:

The diagnosis of the troubles of the idol-producing mind is part of a project of offering “helps” to the human faculties. In several places, Bacon renders this idea by means of medical language: both the investigation of the sources of the idols and the method for the legitimate pursuit of natural knowledge are invested with the role of a “cure” or “remedy” or “purging” of the mind . . . There is thus a mental-medicinal aspect to Bacon’s epistemological and methodological project that allows us to extend the relevance of his phrase *medicina mentis*, used explicitly in the context of his moral philosophy, to his reflections on the best way to proceed in natural philosophy.

These remarks appear on p. 184 of Sorana Corneanu and Koen Vermeir, “Idols of the Imagination: Francis Bacon on the Imagination and the Medicine of the Mind,” *Perspectives on Science* 20, no. 2 (2012): 183–206. For an authoritative treatment of this whole tradition surrounding the Royal Society, see Sorana Corneanu, *Regimens of the Mind: Boyle, Locke, and the Early Modern Cultura Animi Tradition* (Chicago: Chicago University Press, 2011).

Several figures helped to domesticate this way of thinking about philosophy in the German tradition, though from largely opposed intellectual standpoints. Johann Joachim Lange (1670–1744) was part of a group of Pietist thinkers who gravitated towards Christian Thomasius in Halle. His *Medicina Mentis*, first published in 1704, offers a history of philosophy in which various kinds of corrupted or diseased thinking, stemming ultimately from the Fall, are critically examined.⁴⁶ Thomasian Pietists, particularly in Saxon lands, eagerly adapted this therapeutic approach to philosophy for their own distinctively spiritual ends.⁴⁷ Fichte was steeped early on in this Pietist tradition, particularly when it came to ideas about educational reform that had first been implemented at Halle and then emulated later at Leipzig and Jena.⁴⁸

The discourse of philosophy as therapy was not, however, limited to Pietist or Thomasian circles. Coming from an opposed standpoint, E. W. von Tschirnhaus’s principal work is entitled *Medicina Mentis, sive Artis Inveniendi praecepta generalia*.⁴⁹ It first appeared in Amsterdam in 1695, and then again, in a more complete form, in Leipzig several years later. A largely forgotten figure who nevertheless played a large role in some of the important philosophical developments of the seventeenth and eighteenth centuries, as well as in the development of Meissen chinaware, Tschirnhaus served as a direct conduit of the Ciceronian tradition into German intellectual life. In his main work, Tschirnhaus advocates an approach to logic as an applied or practical discipline, very much in the vein of a larger neo-Stoic movement that was taken up by Descartes and Spinoza. He emphasizes the importance of discovering the truth on one’s own and for oneself, as opposed to more scholastic models of doing

⁴⁶ For a discussion of Lange, see Martin Mulsow, “Eclecticism or Skepticism? A Problem of the Early Enlightenment,” *Journal of the History of Ideas* 58, no. 3 (1997): 465–77.

⁴⁷ See Kelly J. Whitmer, “Eclecticism and the Technologies of Discernment in Pietist Pedagogy,” *Journal of the History of Ideas* 70, no. 4 (2009): 545–67.

⁴⁸ For a good general discussion of these Thomasian ideas about educational reform, see Thomas Ahnert, *Religion and the Origins of the German Enlightenment: Faith and the Reform of Learning in the Thought of Christian Thomasius* (Rochester: University of Rochester Press, 2006).

⁴⁹ My discussion of Tschirnhaus here is indebted to C. A. van Peursen, “E. W. von Tschirnhaus and the *Ars Inveniendi*,” *Journal of the History of Ideas* 54, no. 3 (1993): 395–410.

philosophy, which he characterizes as *philosophia verbalis*. It is useful to recall here Fichte's emphasis, at the conclusion of the "Facts of Consciousness" lectures, that one must "actually have" the thought of knowing as an independent life. Tschirnhaus also advocates pedagogical reforms that certainly resonate with what Pestalozzi and, later, Fichte, similarly propounded. His ideas about pedagogical reform indeed had some influence on eighteenth-century discussions. Tschirnhaus's ideas were widely received, influencing Wolff and thus entering into the philosophical culture in which Fichte's own thinking emerged. One interesting connection is Tschirnhaus's suggestion that the static term "idea" should be replaced by the verb "to conceive" (*concipiri*). In Fichte's time, this rationalist tradition had recently been revived by C. G. Bardili, a philosophical ally of K. L. Reinhold whose work Fichte read closely. Bardili's critique of the Kantians (including Fichte) can be found in his 1800 *Grundriss der Ersten Logik*, the subtitle of which indicates that the work is "not a critique but rather a *medicina mentis*, chiefly useful for Germany's critical philosophy."

The lectures on the "Facts of Consciousness" can be profitably read as a Fichtean (and thus, ultimately, transcendental) iteration of this traditional idea. What the lectures aim to establish is the properly transcendental standpoint as that from which any "scientific" philosophy must proceed. Fichte's particularly radical conception of that standpoint means that, by his lights, not even Kant had succeeded in fully adopting it, and so had failed to execute completely the task of making philosophy scientific.⁵⁰ More pressingly, the failure to appreciate the standpoint

⁵⁰ It is important to mention here that not everyone agrees that this goal of making transcendental philosophy "scientific" was either necessary or desirable. Karl Ameriks, for example, argues forcefully that the combination of foundationalism (entailed, on his view, by this "scientific" revision of Kant) and of a radical version of the "primacy of practical reason" thesis violated not only the *letter* of Kant's thought (as Fichte admitted), but also its *spirit*. Ameriks makes this point in a series of essays, as well as in a monograph: "Kant, Fichte, and Short Arguments to Idealism," *Archiv für Geschichte der Philosophie* 72, no. 1 (1990): 63–85; "Fichte's Appeal Today: The Hidden Primacy of the Practical," in *The Emergence of German Idealism*, ed. Michael Baur and Daniel O. Dahlstrom (Washington, DC: Catholic University of America Press, 1999), 116–30; "The Practical Foundation of Philosophy in Kant, Fichte, and After," in *The Reception of Kant's Critical Philosophy: Fichte, Schelling, and Hegel*, ed. Sally Sedgwick (Cambridge: Cambridge University Press, 2000), 109–28; and *Kant and the Fate of Autonomy: Problems in the Appropriation of the Critical Philosophy* (Cambridge: Cambridge University Press, 2000).

from which transcendental philosophy proceeds paved the way for the misunderstanding of Fichte’s idealism as perniciously subjective. What the lectures aim to achieve, then, is not so much to prove or demonstrate a conclusion, but rather to bring Fichte’s listeners to adopt freely for themselves the required standpoint. For this reason, the lectures read as a kind of iterated logical psychoanalysis, in which the audience is repeatedly directed to engage in a kind of recursive action, to “peer into your own consciousness” and to engage in a sort of “experiment” on yourself.⁵¹

This process involves the audience actually enacting the gradual “liberation” of “knowing” from its initial state of “boundedness,” as Fichte puts it. Beginning first with sensible or external perception, Fichte tries again and again to coax out the intuition that “consciousness is not merely a dead and passive mirror of external objects,” but is “itself something living and dynamic.”⁵² An alternative theory of perception may, naturally, be achieved along the way, yet the primary aim is the actual intuition of the mind’s active nature. Following his discussion of external perception, Fichte takes up both reflection (in the Lockean sense of self-awareness) and various forms of imaginative abstraction. That both are, as he says, acts of “freedom”⁵³ is made apparent by contrasting the supposedly constrained nature of a child’s encounter with the world and the imaginative freedom of a scientist (for example, Archimedes).⁵⁴ Part of what Fichte aims to achieve here is the intuition of the famously Fichtean point that theorizing is fundamentally *practical*, or that science is a product of freedom. But, as the difference between the imaginative freedom of the scientist and the unrestrained delusions of the lunatic indicates,⁵⁵ this activity of freedom remains

Ameriks’s position has itself been challenged, for example by Daniel Breazeale, in “Two Cheers for Post-Kantianism: A Response to Karl Ameriks,” *Inquiry* 46, no. 2 (2003): 239–59.

⁵¹ StA 1:231.

⁵² StA 1:235.

⁵³ StA 1:242.

⁵⁴ StA 1:243–4, 248.

⁵⁵ StA 1:248.

somehow incomplete unless the “practical faculty” is guided by “purposive concepts” (*Zweckbegriffe*).⁵⁶

As Fichte engages the audience in the further analysis of purposive action, he comes closer and closer to the goal of the actual thought “that knowing is absolutely [*schlechtweg*] a self-sufficient life.” Purposive action turns out to require a world shared with “creatures like me outside of me,”⁵⁷ and of objects or perceivable events that are products not simply of nature but of rational agency. The latter, in turn, requires a notion of some sort of non-natural restraint on action, or, in other words, a *moral law*.⁵⁸ The moral law is the expression of what Fichte calls the “moral nexus” (*Zusammenhang*), a totality or “world” of individuals, each of whom one is constrained to recognize as being an agent, rather than simply a force of nature.⁵⁹ As Fichte puts it, what he is trying to bring into view is a sort of incapacity that is not merely a *physical* incapacity, but rather is “a different, *moral consciousness* of a *may not* [*Nichtdürfens*].”⁶⁰ Recall here his comments on the difference between the imaginative freedom of a scientist and the unrestrained imagination of a lunatic: the “higher capacity” for constraining one’s activity through a concept had already been operative at that level, but is only fully apparent or articulate in the moral domain.

Fichte cashes out the difference between a “moral nexus” and a purely physical one in these terms: assuming a plenum, the locomotion of one object “is *simply* one and the same as that of *something else*, and there is no middle term between them.”⁶¹ In a “moral nexus,” on the other hand, “there is something that mediates the self-determination of the cause and the determination of the effect—this must be a self-determination of the cause in the other.”⁶² This

⁵⁶ StA 1:278.

⁵⁷ StA 1:294–7.

⁵⁸ StA 1:331.

⁵⁹ StA 1:338–48.

⁶⁰ StA 1:353.

⁶¹ StA 1:336.

⁶² StA 1:337.

observation is supposed to lead to the following thought: “the one person acts [*wirkt*]; this is a self-determination and, as such, it remains completely within him. But immediately united to this self-determination there arises a thoroughly universal consciousness for all that immediately brings along with it a limiting *ought*: in this way, a moral connection is achieved between all.”⁶³ While subjective self-awareness is only possible in an individual, the kind of mediated moral awareness described above must rest on some underlying unity of life that is at once efficacious and not reducible to individual subjects.⁶⁴ As Fichte’s discussion proceeds, then, we are supposed to move beyond the intuition of our own freedom to that of the dynamism, vitality, and freedom of the whole in which each individual participates and which is expressed in the self-consciousness of individual agents.

Fichte then asks, to what end this “drama” (*Schauspiel*) of expressions of the life of the totality? His answer is that “the one life of freedom is at bottom nothing but the form of the intuition of morality.”⁶⁵ The idea that the sheer expression of life-force cannot be an end in itself is something Fichte grounds on further facts of consciousness, including our natural resistance to the idea.⁶⁶ The claim here seems to be that, if we just accept at this point that a naturalistic view of things is ultimate, then we are failing to do justice to a noteworthy “fact of consciousness”—namely, moral awareness.⁶⁷ Clearly, there is nothing approaching deductive necessity here. The latter, as Fichte makes plain in comments I have already referenced, is not to be found in the “experience” (*Empirie*)

⁶³ StA 1:337.

⁶⁴ StA 1:348.

⁶⁵ StA 1:358.

⁶⁶ StA 1:359.

⁶⁷ Cf. Fichte’s remark from a set of notes composed in 1807 (GA 2/10:54), when he was working on a reply to Jacobi’s charge of “nihilism,” leveled in the latter’s 1799 open letter: “Very apt [*treffende*] formulation against Jacobi. The true holy place is a *life* (i.e., the moral [life]), that of the concepts that are the *exponents* of a life, simply as such. This is what I should [be taken to have] said with the moral world order; not a dead, but rather a living God.”

of the “facts of consciousness,” but rather in the *Wissenschaftslehre* proper. Instead, what Fichte is driving towards is a perspective in which both physical nature and individuality lose their semblance of “autonomy” (*Autonomie*),⁶⁸ for the sake of a kind of eternal point of view or vision that transcends and unifies the world of becoming, that is, for the sake of becoming the “image of God.”⁶⁹ Religion, Fichte says, “completes [*vollendet*] the life of knowing, and is its highest peak.”⁷⁰

The preceding account of the course of the lectures is necessarily brief and leaves out a number of details. My hope is that Fichte’s argumentative strategy is nonetheless fairly clear. Starting from the “boundedness” of life in the external perception of particulars, Fichte tries to draw forth the intuition of the mind’s activity. The distinctive character of the mind’s activity, however, becomes most apparent in the moral sphere, which is itself only fully intelligible on the basis of the further intuition of the interconnected totality of “knowing.” This totality is not reducible to a collection of particulars, and thus intuiting it requires that one move beyond the “historical intellect” that characterizes the “natural consciousness.” From the standpoint of the whole, the necessity of the “forms of knowing” can be properly deduced for the first time. Fichte’s claim is that, since it is the standpoint of the whole, the idea that the “forms of knowing” are simply contingent features of an individual’s subjective constitution loses its traction. The emergence of the “forms of knowing” can thus be seen as *epigenetic*, as the production of individual forms or structures from out of a self-contained totality. With this, Fichte claims that the legacy of Kant’s transcendental turn has arrived at its genuine fulfillment.⁷¹

⁶⁸ StA 1:365

⁶⁹ StA 1:385, 388.

⁷⁰ StA 1:391.

⁷¹ A very early draft of this chapter was presented at the 2012 meeting of the North American Fichte Society in Quebec. Since that time, I have also received valuable comments from Owen Ware, Elijah Millgram, and participants in the 2015 Conference on Transcendental Philosophy and Metaphysics in Osaka, Japan. Matt Haber helped me navigate the history of biology.

Benjamin D. Crowe teaches at Boston University and the University of Utah. He specializes in modern German philosophy, especially Heidegger, Fichte, early romanticism, and Dilthey. He is the author of *Heidegger's Phenomenology of Religion: Realism and Cultural Criticism* (Indiana University Press, 2007) and *Heidegger's Religious Origins: Destruction and Authenticity* (Indiana University Press, 2006). More recently, he has edited *The Nineteenth Century Philosophy Reader* (Routledge, 2015) and has published a translation of J. G. Fichte's *Lectures on the Theory of Ethics (1812)* (SUNY Press, 2016).

6

From Transcendental Philosophy to Hegel's Developmental Method

William F. Bristow

Recently it has been common to read Hegel as employing a version of Kant's transcendental method of argumentation in his *Phenomenology of Spirit*.¹ This is perhaps surprising, given that Hegel in many texts

¹ Famously and influentially, Charles Taylor interprets the opening arguments of the *Phenomenology* as transcendental arguments. See his "The Opening Arguments of the *Phenomenology*," in *Hegel: A Collection of Critical Essays*, ed. Alasdair MacIntyre (Notre Dame: University of Notre Dame Press, 1972), 157–87. Robert Pippin's impactful reading of Hegel as aiming to complete Kant's philosophical project ascribes to Hegel a transcendental method of argumentation. See his *Hegel's Idealism: The Satisfactions of Self-Consciousness* (Cambridge: Cambridge University Press, 1989). Robert Stern gives a nuanced and qualified reading of Hegel's method as employing transcendental argument against a form of skepticism in "Hegel, Scepticism and Transcendental Arguments," in *Skeptizismus und spekulatives Denken in der Philosophie Hegels*, ed. Rolf-Peter Horstmann and Hans Friedrich Fulda (Stuttgart: Klett-Cotta, 1996), 206–25. Frederick Neuhauser extends Taylor's construal of Hegel as following a transcendental method in the "Consciousness" chapters of the *Phenomenology* into his transition to self-consciousness as desire in "Deducing Desire and Recognition in the *Phenomenology of Spirit*," *Journal of the History of Philosophy* 24, no. 2 (1986): 243–62. Rolf-Peter Horstmann, responding to some criticisms of

W.F. Bristow (✉)

Department of Philosophy, University of Wisconsin, Milwaukee, USA

criticizes Kant's philosophical procedure and develops his own in opposition to the latter's. Hegel's procedure in his *Phenomenology of Spirit* can be productively interpreted as a transformation of Kant's critical procedure,² but readings of Hegel's method in that work as "transcendental," in imitation of Kant, miss the respects in which he rejected and aimed to improve on Kant's Critical Philosophy. Hegel's phenomenological method is one of the most intriguing parts of his philosophy, but it can only be appreciated to the extent that the method is not interpreted so as to conform to a Kantian paradigm that Hegel rejected.

Kant's Transcendental Method

Readers of Hegel can attribute to him Kant's transcendental method, despite his explicit criticisms and rejection of Kant's method, because "Kant's method" and "the transcendental method" are sufficiently ambiguous. It is useful, when addressing the question of the relation of Hegel's method to Kant's, to distinguish "the transcendental argument form" from Kant's *critical* method. The discussion of transcendental arguments focuses on an *argument form* classically exemplified in key arguments in Kant's *Critique of Pure Reason* (in particular, the transcendental deduction of the categories, the Second Analogy of Experience, and the Refutation of Idealism). Kant himself presents his work as innovative in method, indeed as "transforming the accepted procedure of metaphysics,"³ but his own presentation of his

interpretations of Hegel's method as transcendental, offers an interpretation of his overall argument in the *Phenomenology* as "transcendentalistic" (a qualified form of transcendental) in "The *Phenomenology of Spirit* as a 'transcendentalistic' argument for a monistic ontology," in *Hegel's "Phenomenology of Spirit": A Critical Guide*, ed. Dean Moyar and Michael Quante (Cambridge: Cambridge University Press, 2008), 43–62.

² I would say that, since I provide a reading of Hegel's project and method in the *Phenomenology* as a self-conscious transformation of Kant's critical project and method in William F. Bristow, *Hegel and the Transformation of Philosophical Critique* (Oxford: Oxford University Press, 2007).

³ Bxxii.

methodological innovation points not to the transcendental argument form, which, having been abstracted from his work, has taken on a life of its own in subsequent philosophy, but rather to *critical* philosophy, in contrast to *dogmatic procedure*.⁴ The question whether Hegel employs Kant's method of transcendental argumentation in his *Phenomenology of Spirit* pulls in its train the question of how the transcendental method of argumentation is related to the critical procedure of philosophy in Kant's work and in Hegel's view of it.

The Transcendental Argument Form

In his classic article, "The Opening Arguments of the *Phenomenology*," Charles Taylor argues that the first three chapters of the *Phenomenology of Spirit*, the section on "Consciousness," are "an essay in transcendental arguments." By "transcendental arguments" he means "arguments that start from some putatively undeniable facet of our experience in order to conclude that this experience must have certain features or be of a certain type, for otherwise this undeniable facet could not be."⁵ Taylor notes that the argument form "is very much part of contemporary philosophical debate," which he illustrates with its prominent employment in two philosophical texts: P. F. Strawson's *Individuals* and Wittgenstein's *Philosophical Investigations*.⁶ The life of this argument form in contemporary philosophy has been very much bound up with its supposed promise as employed against the position of the skeptic regarding empirical knowledge. As an antiskeptical strategy, the transcendental argument *begins from* some claims or phenomena either explicitly shared between the skeptic and opponent of skepticism or, in Taylor's phrase, "putatively undeniable," and from there argues *to* the disputed knowledge or phenomena as *necessary conditions of the possibility* of the shared starting point. In P. F. Strawson's rendering of Kant's key arguments in

⁴ Bxxxv.

⁵ Taylor, "Opening Arguments," 151. He then writes that the arguments are called "transcendental" because the best-known examples of such arguments are to be found in Kant.

⁶ Taylor, "Opening Arguments," 151–8.

his *Bounds of Sense*, he begins from some self-consciousness that even the skeptic does not or cannot deny and then argues that that self-consciousness itself presupposes, or has among its necessary conditions, knowledge of empirical objects.⁷

Taylor interprets Hegel's opening arguments as directed, not against skepticism regarding empirical knowledge primarily, but against something that underlies the threat of modern skepticism, namely, "the epistemological tradition" of modern philosophy and the view of experience in that tradition. In this tradition, experience is "the passive reception of sense-data," conceived as subjective and private.⁸ A central epistemological problem in this tradition is how to justify or explain the possibility of empirical knowledge of objects on the basis of experience, so conceived. Taylor describes the model as "contemplative," based on the fact that the sense-data are passively received and on the fact that the model consequently problematizes causality (classically, and importantly for Kant, in Hume's work). Taylor claims that, although Kant's transcendental arguments begin a tradition of attack on this model,⁹ some of the subsequent thinkers in this tradition (for example, Hegel, Wittgenstein, Heidegger, Merleau-Ponty) have shown that Kant's own epistemology remains significantly shaped by the model. Taylor reads Hegel's opening arguments in the "Consciousness" section of the *Phenomenology of Spirit* as culminating in a transformed conception of human experience according to which our bodily interaction with physical things in our environment is fundamental. Hegel's transcendental arguments in the "Consciousness" section, as interpreted by Taylor, culminate in a conception of our condition that can be characterized using Heidegger's phrase as "being-in-the-world."¹⁰ Although Taylor does not see external-world skepticism as the prime target of the transcendental arguments he reconstructs, he does take arguments of this

⁷ Strawson sums up a main stretch of Kant's transcendental deduction in the following claim: "unity of diverse experiences in a single consciousness requires experience of objects." P. F. Strawson, *The Bounds of Sense: An Essay on Kant's "Critique of Pure Reason"* (London: Methuen, 1966), 98.

⁸ Taylor, "Opening Arguments," 176.

⁹ Taylor, "Opening Arguments," 156–7.

¹⁰ Taylor, "Opening Arguments," 185.

form to begin from a “putatively undeniable facet of our experience,” at least implicitly affirmed by the opponent of the position the argument supports. He refers to this as the “rock bottom starting point” of the argument.¹¹

Kant's Presentation of His Methodological Innovation

As Kant himself presents his innovation in philosophical method, the starting point is not something putatively undeniable or shared between him and his opponents, but rather the condition of metaphysics as “a battlefield of endless controversies.”¹² Underlying this condition of metaphysics, on Kant's diagnosis, is the circumstance that *there is widespread disagreement on the criteria* for rational knowledge, a condition which itself reflects the fact that no one hitherto has clearly posed the question of *how rational knowledge is possible for us at all*.¹³ Though the transcendental method is often presented as answering the question of how some phenomenon is possible, thinly enough described for the description to be shared among all parties in the dispute, it is striking that the “How possible?” question Kant foregrounds in the presentation of his innovation in method is “How *is metaphysics* (as a science) possible?” (assuming that metaphysics as a science does not yet exist). Kant's analysis of this question yields the more specific determination of the question as that of how synthetic a priori knowledge is possible for us, where that question encompasses not only the problematic science of metaphysics, but also the actual sciences of mathematics and pure natural science. Still, Kant foregrounds his revolution in method as a prior inquiry into the possibility of metaphysics, prior relative to the science of metaphysics itself, and as a subjective inquiry, in the sense that

¹¹ Taylor, “Opening Arguments,” 153.

¹² Aviii, Bxiv ff.

¹³ In his *Prolegomena*, Kant claims that reason lacks in its metaphysical researches what he calls “a standard weight and measure” (*ein sicheres Maaß und Gewicht*) by which to distinguish soundness from shallow talk. AK 4:256; *Prolegomena to Any Future Metaphysics*, trans. and ed. J. W. Ellington (Indianapolis: Hackett, 1977), 2.

the question of the possibility of metaphysics is investigated through reflection on our own cognitive capacities.

Kant discovers that the *solution* to the problem of how rational (or, anyway, a priori) knowledge is possible, conceived as requiring the application of a priori concepts to *objects*, requires asking *another* “How possible?” question: namely, “How can *sensible* representations (as subjective modifications of the mind) relate to objects or amount to knowledge of objects?” Otherwise put: “How is empirical knowledge of objects possible for us?” In this way, Kant’s pursuit of his critical method leads him to develop the transcendental arguments that have had such influence.

Within Kant studies, P. F. Strawson’s reading of Kant’s transcendental deduction of the categories as directed against the external-world skeptic has been strongly challenged.¹⁴ In the transcendental deduction, Kant argues that the pure concepts of the understanding have objective validity as necessary conditions of possible experience, where experience, as empirical knowledge of objects, is assumed. Taking empirical knowledge for granted, Kant argues that the application of the a priori concepts in question to objects is *a necessary condition* of the possibility of experience, insofar as synthesis of sensible representations according to the a priori concepts constitutes their relation to objects. Kant’s transcendental deduction of the categories may be seen as exhibiting the form of a transcendental argument, insofar as he argues there that categories apply to appearances as necessary conditions of the possibility of experience, but, on this construal, it is not plausible that the starting point of the argument (experience, empirical knowledge) is undeniable, or that it is implicitly accepted by the skeptic.

The content of both Kant’s starting point and of his conclusion in the transcendental deduction of the categories, and the precise statement of his argument from the former to the latter, are controversial and much-debated. What matters for my purposes and in relation to the question

¹⁴ Particularly worth mentioning here are Karl Ameriks, “Kant’s Transcendental Deduction as a Regressive Argument,” *Kant-Studien* 69 (1978): 273–87; and Stephen Engstrom, “The Transcendental Deduction and Skepticism,” *Journal of the History of Philosophy* 32, no. 3 (1994): 359–80.

of whether Hegel adopts Kant's method is that Kant's transcendental arguments are advanced in the context of a larger critical project that is supposed to effect a revolution in method. According to this revolution, in an effort to establish metaphysics as a science for the first time, in a condition in which the criteria and possibility of rational knowledge are in question, we undertake a reflection on our cognitive faculties to determine how and whether metaphysics is possible for us.

Hegel's Critique of Kant's Method

Hegel articulates the method of his *Phenomenology of Spirit* in the last several paragraphs of the Introduction to that work. He begins the Introduction by criticizing the temptation to proceed in this sort of epistemological inquiry in the way Kant proceeds, namely, by stepping back and asking first, through reflection on our cognitive capacities, how metaphysics (or, as Hegel puts it, "the actual cognition of what truly is") is possible for us.¹⁵ While Hegel suggests that this procedure seems a natural response to uneasiness regarding the possibility of metaphysical knowledge or, as he puts it, the possibility "of securing for consciousness through cognition what exists in itself," the procedure turns out to be self-defeating.¹⁶ The procedure is bound to lead to the conviction, Hegel writes, that "there is a boundary between cognition and the Absolute that completely separates them."¹⁷ Hegel suggests in the opening of the Introduction to the *Phenomenology* that the negative outcome to which Kant's prior inquiry into the possibility of metaphysics leads, namely, the outcome that cognition of what exists in itself is impossible for us, is implicit in Kant's methodological

¹⁵ Hegel's practice in his *Phenomenology* is to refrain from naming names, and there is a good basis for believing that he means to allude in the opening of the Introduction to a general epistemological procedure of which Kant's critical philosophy is but one central exemplification (Locke's in the *Essay* being another). As long as Kant's critical procedure is central among those to which he alludes, that he has others in mind as well does not matter to my purposes.

¹⁶ PS ¶73, PG 68.

¹⁷ PS ¶73, PG 68.

innovation. Indeed, Kant's methodological innovation *presupposes* this negative outcome and, moreover, presupposes the result that our knowledge is restricted to objects as they are *for us* (or, in other words, Kant's method *presupposes* the subjectivism of his idealism). Hegel writes: "above all, it [this epistemological procedure] presupposes that the Absolute stands on one side and cognition on the other, independent and separated from it and yet something real; or in other words, it presupposes that cognition which, since it is excluded from the Absolute, is surely outside of the truth as well, is nevertheless true."¹⁸

Hegel no more than merely insinuates in the Introduction to the *Phenomenology* that Kant's method illegitimately presupposes the skepticism (with respect to absolute knowledge) and the subjectivism of his philosophical system. Hegel says more in other works in criticism of Kant's method, but nowhere does he develop a fully elaborated case against it. Elsewhere I have attempted to develop what I consider Hegel's case, as implied in his various remarks on Kant's method.¹⁹ Here I can only report on Hegel's case, as I understand it, as background to remarks on his "developmental method" which is meant to show it as resolutely opposed to Kant's transcendental method.

As noted earlier, Kant's inquiry takes as its starting point a condition in which agreement on criteria regarding rational knowledge or metaphysics is lacking. The *outcome* of Kant's inquiry is his famous Copernican Revolution in epistemology, according to which "objects must conform to our cognition" rather than our cognition to objects, as supposed hitherto.²⁰ In particular, Kant finds that our cognitive faculties contain a priori "forms" to which the sensible content of our knowledge must conform in order for the objects of our knowledge to be possible for us at all. The highest form to which sensible content must conform, according to Kant's account in the transcendental deduction, is the synthetic unity of apperception. And thus, Kant says, "the synthetic unity of apperception is the highest point to which one must affix all use of the understanding, even the whole of logic and,

¹⁸ PS ¶74, PG 70.

¹⁹ Bristow, *Hegel and the Transformation*, Part I.

²⁰ Bxvi ff.

after it, transcendental philosophy; indeed this faculty is the understanding itself.”²¹ As Hegel reads Kant, the latter essentially advances as the fundamental criterion of human knowledge “conformity of content to *our* understanding” where “our” understanding is conceived of as merely ours, as finite, limited. The subjectivism of Kant’s idealism (that we can know things only as they are for us, not as they are in themselves) is implied in this criterion, as he himself goes on to make explicit. But what Kant does not make explicit, of course, is that his *methodological innovation* already implies this criterion. According to his methodological innovation, we make the epistemological demand that the possibility of rational knowledge be made intelligible to us in a prior subjective reflection, in advance of the science of metaphysics itself, as itself a condition of the possibility of metaphysics as a science. Implicit in this epistemological demand and procedure is the criterion of *conformity of content to us*, to our self-reflection, as finite self-consciousness. This conception of the shape of knowledge precludes (for us) knowledge of what exists “in itself.” In *Faith and Knowledge*, Hegel claims that the philosophy of Kant (together with the philosophies of Jacobi and Fichte) “raised the standpoint of the subject, the standpoint of absolutely existing finitude, to the first and highest place.”²² Hegel emphasizes that this standpoint of the finite “I” is *fixed* (self-standing) over against what is regarded as its other or opposite, that is, the standpoint of truly rational or infinite cognition.

Hegel's Developmental Method

Hegel's method in his *Phenomenology* is determined, to a significant extent, by the ambition to avoid the subjectivism of Kant's idealism. Hegel defines his own phenomenological method at the end of the Introduction to the *Phenomenology*. I will discuss three key elements of that method, with attention to the relation of the method, defined by these elements, to Kant's method.

²¹ B133–4 n.

²² GW 297f.

Criterion-less Critique

Hegel motivates his method with a problem concerning the criterion or standard of epistemological inquiry. He notes that, as the “examination and testing of the reality of cognition” (*Untersuchung und Prüfung der Realität des Erkenntnis*), the inquiry seems not to be able to take place without some sort of presupposition, which can serve as its underlying criterion (*Maßstab*). The problem is, in *this* inquiry, nothing “has yet justified itself” as the criterion of cognition.²³ What *counts as* rational knowledge, if anything does, is just what is primarily in question in this inquiry. Given this, the problem is: how can the inquiry (into “the reality of cognition”) so much as begin?

Hegel strives to avoid two opposed dogmatic ways of proceeding, two opposed ways of “presupposing a criterion” in the testing of the reality of cognition. In light of the allusions to Kant’s critical procedure in the opening of the Introduction, Hegel clearly means to avoid presupposing “conformity to ourselves,” conformity to the standpoint of the finite reflective self-consciousness, as the ultimate standard of knowledge. That is, he wants to avoid presupposing in the investigation Kant’s “Copernican Revolution in epistemology,” according to which objects must conform to us, with the consequence that knowledge of what exists *in itself* is impossible for us. But he also wants to avoid the way associated with Schelling of presupposing that “we” already possess a justified criterion of metaphysical knowledge, independently of this inquiry, a criterion that we can simply apply in the examination or testing of other philosophical standpoints as they come before us.

Hegel claims that this problem of the criterion is overcome by “the nature of the object we are investigating.”²⁴ “Consciousness provides its criterion from within itself,” he claims, “so that the investigation becomes a comparison within it.” That is, the knowing subject distinguishes between (a) what is to it true and (b) the basis on which it takes something to be true. “In consciousness,” Hegel writes, “one thing exists

²³ PS ¶81, PG 75–6.

²⁴ PS ¶84, PG 76.

for another, i.e., consciousness regularly contains the determinateness of the moment of knowledge; at the same time, this other is to consciousness not merely *for it*, but is also outside of this relationship, or exists *in itself*: the moment of truth.”²⁵ In other words, consciousness makes knowledge claims (something is to consciousness true), but consciousness also has a conception of that in virtue of which its claims are allegedly true. Hence the knowing subject (consciousness) has its own criteria to apply in assessing its own knowledge claims. The general point is that, because of the duality within consciousness, there are two elements that can be compared in testing knowledge, without our importing criteria of our own.

I call attention here to how Hegel's phenomenological method as the immanent critique of consciousness clashes starkly with the transcendental form of argumentation, as defined above. A transcendental argument *proceeds from* some common ground among disputants, something not in dispute in whatever the controversy, and then proceeds to allegedly necessary conditions of this shared common ground, which *are* the focus of the controversy among the disputants. Hegel's method, in contrast, is explicitly designed to proceed in the condition in which *there is no recognized common ground*, a condition in which the disputants or inquirers cannot be supposed to share in judgments or criteria at all. His method is designed to proceed in a context in which nothing is taken as fixed or established. That the scene of this inquiry is one in which the disputants share *no common ground* is reinforced in the Preface, where Hegel represents the starting point of this inquiry as one in which the “standpoint of consciousness which knows objects in their antithesis to itself, and itself in antithesis to them, is for Science the antithesis of its own standpoint.”²⁶ Each is to the other, he says, “the inversion of truth.” If consciousness's conception and science's conception of knowledge are to “come to terms,” given that they are diametrically opposed, then we need a method that proceeds without taking anything as fixed or agreed upon, in stark contrast to the transcendental

²⁵ PS ¶184, PG 77.

²⁶ PS ¶126, PG 30.

form of argumentation. And it is just such a method that Hegel defines in the Introduction.

In his article, Taylor discusses Hegel's presentation of his method in the Introduction. He notes that Hegel's aim in the *Phenomenology* "is to move from the 'natural,' i.e. commonsense, view of consciousness to his own."²⁷ Taylor recognizes that Hegel wants his criticism of consciousness "to take nothing for granted." Taylor recognizes further that Hegel exploits the fact that "a notion of experience contains its own 'yardstick,'" that is, its own idea of what experience is, and thus can be criticized according to its own criterion.²⁸ Hegel's procedure consists, on Taylor's interpretation, in consciousness testing its idea of what experience is against its attempts to realize its model in actual claims to know (what Taylor calls "effective experience"). Taylor writes: "if it turns out that effective experience guided by the model contradicts it . . . then [the model] will be shown to be impossible and will have to be changed."²⁹

Taylor claims that Hegel's procedure presupposes "that we can characterize effective experience in terms independent of the model we are working with."

Moreover, if we are to show that the model is not just unrealized in a given case, but cannot be realized, we have to be able to identify some basic and pervasive facets of experience independently of our model (they must be independent, i.e., not derivable from the model itself, if they are to contradict it and show it to be impossible). Hence the method that Hegel outlines in his Introduction to the *Phenomenology* can only be applied if such basic facets can be picked out, and his arguments will stand only to the extent that they can be shown as beyond question.³⁰

The "basic and pervasive facets of experience," identified independently of any model or idea of experience that consciousness harbors, are the

²⁷ Taylor, "Opening Arguments," 157.

²⁸ Taylor, "Opening Arguments," 158.

²⁹ Taylor, "Opening Arguments," 159.

³⁰ Taylor, "Opening Arguments," 160.

starting point(s) of the transcendental arguments Hegel advances, on Taylor's reading the "rock bottom" starting point that a specifically transcendental argument requires, from which such an argument proceeds to articulate the necessary conditions of their possibility.

In presenting his method, Hegel makes no mention of such independently identified facets of experience or knowledge as required by his method, much less of a "putatively undeniable starting point" that is to be agreed to by all parties throughout the inquiry. On the contrary, Hegel's effort, on the face of it, is to devise a method that presupposes nothing as already fixed or established or agreed upon by all parties to the inquiry, as a starting point or first premise of the inquiry itself. Moreover, Taylor's brief argument to the effect that Hegel's procedure presupposes this element is not convincing. One must grant that, if consciousness were simply to mold its knowledge claims—its "effective experience," in Taylor's terms—expressly to fit the model of experience or knowledge that it means to test, no real test would result. But it does not follow from that point that the method presupposes a putatively undeniable starting point that is held fixed throughout the inquiry. The method requires only that the two elements compared in the testing are sufficiently independent of each other that, when consciousness compares them and finds that they do not agree, it learns of the inadequacy of its model.³¹ How consciousness learns this is reflected in the second feature of Hegel's method.

³¹ Michael N. Forster provides a sharp, detailed critique of Taylor's reading of Hegel's method in *Hegel's Idea of a "Phenomenology of Spirit"* (Chicago: University of Chicago Press, 1998), 161–5, especially n. 87. Stephen Houlgate also criticizes Taylor's reading of Hegel's method, specifically for attributing to Hegel the position that the critique proceeds from a *philosopher's* presupposition. Because the presupposition is not shared by the proponent of the configuration of consciousness criticized in relation to it, the resulting refutation, on Taylor's reading, and contra Hegel's purposes, does not amount to a "self-refutation." See Stephen Houlgate, "Is Hegel's *Phenomenology of Spirit* an Essay in Transcendental Argument?" in *The Transcendental Turn*, ed. Sebastian Gardner and Matthew Grist (Oxford: Oxford University Press, 2015), 173–94. In an article responding to Houlgate, Robert Stern argues that, although Houlgate has a point against Taylor's attribution to Hegel *transcendental arguments*, given Hegel's effort to proceed in his inquiry without presuppositions, he continues to argue by employing *transcendental claims*. See Robert Stern, "Taylor, Transcendental Arguments, and Hegel on Consciousness," *Hegel Bulletin* 34, no. 1 (2013): 79–97.

The Transformation of Consciousness

Hegel's desire that the method reflect "fluidity" rather than "fixity" is accentuated in the second feature of his method, as defined in the Introduction, namely, that it results in the transformation (*die Umgestaltung*) of consciousness.

Hegel writes that if the two elements in the formation of consciousness fail to correspond to one another in the testing, then *both* its knowledge claims *and* its criterion change in concert with each other: "the criterion is altered when that for which it was to have been the criterion fails to pass the test; and the testing is not only a testing of what we know, but also a testing of the criterion of what knowing is."³² The elements of the configuration of consciousness (*die Gestalt des Bewußtseins*) form a whole, determined *partly* in relation to each other, and when they prove to be untrue, through failing the test of being in agreement with each other, the whole configuration must change. To suppose that we could proceed by fixing one and making the other conform to it would presuppose that we have independent access to the truth. In the absence of that independent access, there emerges a *new configuration* of consciousness, with a new object and a new criterion.

Taylor speaks of the "transformation" of consciousness when its test fails,³³ but what he means by this transformation, it seems, is simply the change in consciousness's model or picture of experience. What is at stake in Hegel's inquiry is not one's experience, on Taylor's reading, but one's "picture" (or notion or model) of experience. Hegel's inquiry achieves its aim when the picture of experience corresponds to "the reality of experience," which is conceived as being what it is independently of one's picture.³⁴ In this respect, the place of experience in relation to the philosophical account, in Taylor's Hegel, corresponds to that of experience in relation to Kant's transcendental philosophy. But Hegel himself seems particular about how to understand "experience" (*Erfahrung*) in his

³² PS ¶85, PG 55.

³³ Taylor, "Opening Arguments," 158–9.

³⁴ Taylor, "Opening Arguments," 184.

story, and it is not, for him, something given and fixed that we want through the inquiry to get a proper conception of. Rather, for Hegel, explicitly, the experience of consciousness is its *dialectical movement* from one formation to the following, in other words, its transformation.³⁵ Consciousness's experience, in Hegel's story, is its fundamental reorientation to the world, which is required by its self-criticism, since its orientation is partly constituted by its more or less implicit *conception* of itself and its relation to the world.

That the succeeding formations of consciousness have their own criteria, as well as their own objects, seems to imply an incommensurability between these standpoints. How can someone occupying the standpoint of sense-certainty, for example, debate fruitfully someone occupying the standpoint of perception, if they do not share criteria or objects, if they do not agree on what *counts as a reason* for what? So far, it seems that this method at best *explains* why metaphysics is in its lamentable condition ("a battlefield of endless controversies," in the absence of shared criteria); it does not seem designed to institute metaphysics as a science. But Hegel insists that there is an element of "scientific comprehension" to this method, and he tells us that this element is "our contribution," which is the third key element of Hegel's method.

"Our Contribution"

As the German word "*Erfahrung*" connotes, consciousness *learns something* in the transition from one formation to the following; this learning is expressed in its having a new object. But consciousness itself *cannot give an account* of what it has learned; it cannot justify its new object. Whereas consciousness itself experiences the arising of a new object for it, *we phenomenologists* see *how* the new object arises through the skeptical negation of the preceding object. Hegel identifies the "*origination*

³⁵ Hegel: "*Inasmuch as the new true object issues from it, this dialectical movement which consciousness exercises on itself and which affects both its knowledge and its object, is precisely what is called experience*" (PS ¶86, PG 78).

of the new object that presents itself to consciousness without its understanding how this happens” with the *necessity* of the progression, by virtue of which this “*way to Science* is itself already *Science*.”³⁶ This constitutes “our contribution” to the proceedings. We make this contribution, not through applying a criterion of our own, but through mere observation, by merely observing how the new object arises through the skeptical negation of the preceding.³⁷

By virtue of “our contribution,” the procedure yields rational justification, but this rational justification has a distinctive form. Rational justification does not proceed by applying an agreed-upon criterion of what counts as a reason. Again, such criteria are exactly what is in question in the inquiry. Instead, rational justification proceeds *developmentally* (or perhaps *narratively*). That is, by virtue of this method, we are able to *give an account* of the criteria and cognitions within a particular configuration of consciousness by *recounting the story* of how they arise through the self-critical path described in the *Phenomenology*.

An example from the *Phenomenology* will help to illustrate this developmental form of justification. When in the course of its progression consciousness overcomes the one-sidedness of consciousness (which takes the object over against itself as “the True”) and the opposite one-sidedness of *self-consciousness* (which takes itself, or consciousness, to be the truth of the object), consciousness attains to the standpoint of reason, which Hegel characterizes as “the certainty of consciousness that it is all reality.”³⁸ For consciousness as reason, reality is not other than itself, and it is not other than reality. Hegel says that reason supposes “that what *is*, or the in-itself, only *is* in so far as it is *for* consciousness, and what is *for* consciousness is also *in itself* or has *intrinsic* being.”³⁹ At first, anyway, consciousness as reason simply assumes this orientation to reality, is simply *certain* of it, and thus its

³⁶ PS ¶¶87–8, PG 79–80.

³⁷ Cf. Kenley Royce Dove, “Hegel’s Phenomenological Method,” *Review of Metaphysics*, 23 (1970): 615–41.

³⁸ PS ¶233, PG 179.

³⁹ PS ¶233, PG 180.

own “justifications” or stretches of reasoning simply presuppose it. Thus, early in *Observing Reason*, reason simply assumes that the system of genera, species, natural kinds that observing reason constructs in exploring and analyzing nature on the basis of observation “is Nature’s own system.”⁴⁰ The question naturally arises: by what *right* does reason presuppose this?⁴¹ If we were to ask consciousness to justify its presupposition, this would amount to a request for a justification of its criterion, of its “certainty that it is all reality.”⁴² Consciousness as reason, like the preceding configurations of consciousness, cannot, at least when it first comes on the scene, justify its own criterion. Hegel explicitly says that consciousness as reason “merely *asserts* that it is all reality, but does not itself comprehend this.”⁴³ But he also indicates that *we can see* both that and how reason’s criterion *is justified*. The justification of reason’s certainty resides in the preceding dialectical path which has generated it, the path which reason itself, when it first comes on the scene, has forgotten. This is a justification that proceeds, not by “adducing reasons,” which itself presupposes some shared understanding of what constitutes a reason, and therefore some criterion, but rather by recounting the immanent critique of consciousness through which the criterion emerges.⁴⁴

My characterization of Hegel’s developmental or dialectical method is meant to make vivid its difference both from the transcendental form of

⁴⁰ PS ¶246, PG 190.

⁴¹ This question presses particularly against the background of Kant’s treatment of the possibility of an empirical science of nature.

⁴² The distinction between “our” standpoint and the standpoint of the configuration of consciousness under review is akin to the distinction between transcendental and empirical standpoints implied in Kant’s method. Within a form of consciousness, justification proceeds in the light of an assumed standard or criterion, the standard or criterion that is partly constitutive of that particular configuration. But from “our” standpoint, the criterion is not fixed or assumed, but exactly what is in question.

⁴³ PS ¶233, PG 180.

⁴⁴ In his discussion of the method in the Preface, Hegel writes that “it is not difficult to see that the way of asserting a proposition, adducing reasons for it, and in the same way refuting its opposite, is not the form in which truth can appear. Truth is its own self-movement” (PS ¶48, PG 47). This follows his claim in the same context that the truth “includes the negative also.”

argumentation and from Kant's method, as he himself describes it. The essential difference is that, whereas the transcendental form of argument begins from something taken as fixed, as unquestioned, and whereas Kant's method implicitly presupposes the standpoint of the finite, criticizing the subject as fixed, Hegel's method is explicitly devised to proceed without taking anything for granted as fixed from the outset. Thus it is implausible, I believe, to interpret Hegel as employing a transcendental method, as derived from Kant.⁴⁵

Hegel's Developmental Method and Hegel's Metaphysics

I end by raising a question about, or a problem with, Hegel's method as I have outlined it here. Does Hegel's method depend on his metaphysics? If so, that would perhaps be disappointing. What seems initially promising and attractive about Hegel's method is that it enables philosophical critique to proceed without prior commitments, in maximal openness regarding the outcome. Moreover, the ever-renewed interest in Kant's transcendental method (interpreted, anyway, as the transcendental argument form) has been underwritten by the fact that philosophers are able to detach it from the context of Kant's own idealistic system of philosophy and make productive use of it in other contexts and toward other ends. Is Hegel's method similarly detachable from the overall metaphysical position Hegel uses it to defend?

Yes and no, it seems. Hegel's method has a *skeptical use* independent of his metaphysics. The method can be used to undermine philosophical positions without the employment of dogmatically imposed external criteria. Moreover, the method can be employed to show how certain positions are the dialectical outcome of the skeptical negation of other positions. As indicated above, the method enables us to give a rational

⁴⁵ Hegel's debt to Kant, in terms of method, is not to the procedure of the Transcendental Analytic in the *Critique of Pure Reason* nearly so much as it is to the dialectical method of the Transcendental Dialectic.

account of the positions generated through the dialectic, without employing criteria of our own. This use of the method does not depend on the fact that the method arrives, in Hegel's employment of it, at the standpoint of his metaphysical system.

However, it seems that we have no good reason to trust in any of the positions or standpoints arrived at through this method, except insofar as they withstand the internal testing to which the previous forms of consciousness are subjected. In fact, given that the previous forms of consciousness all succumb to skeptical negation through the test that this method applies, perhaps we (or any practitioner of this method) have good reason to incline toward skepticism unless or until the method leads us to "the point where knowledge no longer needs to go beyond itself," as Hegel puts it in the Introduction, "where knowledge finds itself, where concept corresponds to object and object to concept."⁴⁶ Hegel himself says explicitly here that short of this goal "no satisfaction can be found at any of the stations on the way." His dialectical method yields *knowledge*, beyond skeptical refutations, only insofar as it comes to an end in a position that cannot be skeptically negated in turn by the employment of the same method. How does this happen?

Hegel's dialectical method produces a sequence of positions, each succeeding position a result of the skeptical negation of the immediately preceding one. It is hard to see how such a method leads anywhere (that is, produces knowledge), unless it leads where it leads in Hegel's employment. In Hegel's employment, the method achieves its end not in an ultimate member of the sequence, but in the consciousness of the *completeness* of the series, that is, in achieving a point where all the moments of the series form a *whole*. In labeling metaphysical reality "spirit" (*Geist*), Hegel means to mark that that substance becomes what it is through a process, among other things, of knowing itself. In the Preface he says of this metaphysical reality that "it is the process of its own becoming, the circle that presupposes its end as its goal, having its end also as its beginning."⁴⁷ The method too, although it is not

⁴⁶ PS ¶180, PG 74.

⁴⁷ PS ¶18, PG 23.

supposed to presuppose anything, achieves a resting point only through describing a self-returning circle, encompassing all significant positions as stages in spirit's self-knowing and self-becoming. It is striking that a method initially motivated as maximally noncommittal and open, turns out to bring in its train such an encompassing and demanding metaphysics. It is not surprising, of course, that Hegel employs the method to arrive at his own metaphysical position. My point here, though, is that it is hard to see how the method, despite Hegel's convincing motivation of it as maximally noncommittal and open, has any non-skeptical application outside of the context of his own metaphysics.⁴⁸

William F. Bristow is Associate Professor of Philosophy at the University of Wisconsin, Milwaukee. His research focuses particularly on German philosophy of the eighteenth and nineteenth centuries. He is the author of *Hegel and the Transformation of Philosophical Critique* (Oxford University Press, 2007).

⁴⁸ In addition to the sources cited above, this chapter is indebted to: Karl Ameriks, *Kant and the Fate of Autonomy* (Cambridge: Cambridge University Press, 2000); Paul Franks, *All or Nothing: Systematicity, Transcendental Arguments, and Skepticism in German Idealism* (Cambridge, MA: Harvard University Press, 2005); Paul Franks, "Transcendental Arguments, Reason and Scepticism: Contemporary Debates and the Origins of Post-Kantianism," in *Transcendental Arguments: Problems and Prospects*, ed. Robert Stern (Oxford: Clarendon Press, 1999), 111–45; Jean Hyppolite, *Genesis and Structure of Hegel's "Phenomenology of Spirit,"* trans. Samuel Cherniak and John Heckman (Evanston: Northwestern University Press, 1974); Angelica Nuzzo, "Dialectic as Logic of Transformative Processes," in *Hegel: New Directions*, ed. Katerina Deligiorgi (Montreal: McGill-Queen's University Press, 2006), 85–103.

7

How Transcendental Is Cohen's Critical Idealism?

Halla Kim

In this chapter, I discuss Hermann Cohen's critical idealism and examine the extent to which it owes to, and goes beyond, Kant's transcendental philosophy. I suggest that, while remaining within the venerable tradition of transcendental philosophy, Cohen goes well beyond Kant and reveals a new dimension of this philosophy. In the first section, I introduce Cohen's conception of the transcendental. In the next section, I discuss how this conception leads to his transcendental method as he adapts the analytic method of Kant's *Prolegomena*. In the third section, I show how and why Cohen cannot accept Kant's synthetic method and its associated psychologism. In the fourth section, I suggest that Cohen completes the transcendental method in a new way, by means of what might be called "the method of hypothesis" and the ensuing doctrine of origin. In the next section, I critically discuss the view that there is no room for a subject in Cohen's critical idealism, despite its transcendental orientation.

H. Kim (✉)

Department of Philosophy and Religion, University of Nebraska at Omaha,
Omaha, USA

e-mail: hallakim@unomaha.edu

© The Author(s) 2016

H. Kim, S. Hoeltzel (eds.), *Transcendental Inquiry*,

DOI 10.1007/978-3-319-40715-9_7

I argue that, even though it practically abandons the Kantian synthetic method in favor of the analytic method followed by the method of hypothesis, Cohen's critical idealism goes well beyond the purview of the Kantian project with his recourse to origin.

Cohen on the "Transcendental"

Kant is something of a grandfather of all philosophies that claim themselves transcendental. Even though the term "transcendental" is not Kant's coinage, the distinctive kind of philosophy he developed, and the unique method he employed, leave no doubt that he is solely responsible for initiating the tradition. For Kant, the term "transcendental" is referred to cognition, which is occupied not so much with objects but rather with our a priori concepts of objects in general. It is not the first-order reflection on the objects of knowledge, but the second-order reflection on the underlying conditions for our knowledge of objects.¹ The system of such concepts, Kant continues, would be called "transcendental philosophy."²

What is transcendental for Kant then begins with experience (or empirical knowledge), yet at the same time it goes beyond experience because it sheds light on the conditions of the experience. What is transcendental is *eo ipso* a priori, but it is not necessarily the other way around. For example, " $7+5=12$ " is a priori but is not transcendental because it does not state a condition of the possibility of experience. For the transcendental must be able to provide the basis for the possibility of other a priori cognition.³ Kant suggests that a critique of pure reason is a propaedeutic to the complete system of principles, because the former is a mere estimation of pure reason, of its sources and boundaries.⁴

¹ Frederick C. Beiser, *The Genesis of Neo-Kantianism: 1796–1880* (Oxford: Oxford University Press), 466, 483–4.

² A11–12/B25.

³ B151.

⁴ A11/B25.

An organon of pure reason would be the sum total of all those principles in accordance with which all pure a priori cognitions can be acquired and actually brought about.⁵ The entire transcendental philosophy, which necessarily precedes all metaphysics, is nothing but the complete solution of the problem propounded here in systematic order and completeness. In the end, for Kant, the transcendental is supposed to establish the possibility of metaphysics.⁶

Cohen retains the Kantian conception of critique conceived as the transcendental investigation of the pure principles of knowledge.⁷ However, he rejects any prior and separate treatment of the sensible conditions of knowledge in the manner of Kant's first Critique, as this would threaten the autonomy of thought.⁸ In his later period, Cohen conceives thought to be thoroughly self-productive as the sole origin of knowledge and its objects. Thought is not a representative or psychological activity impacted on by the Other, but self-productive, that is, productive in and of itself. In Cohen's view, a critique then must be a transcendental logic without the aesthetic.⁹ Space and time are thus resolved into the categories of thought, that is, in mathematical thought. But even though thought does not need a *given* in sensibility, it does not dispense with what the best of the sciences delivers, that is, the fact (*Faktum*) of science. Rather, thought is supposed to ground its possibility in the principles of pure knowledge.

Cohen makes clear that the hallmark of a philosophical critique lies in its method, which he calls transcendental. For Cohen, a transcendental method consists in proceeding from the hard fact of science in a given branch of human cultural activities and then ferreting out its necessary a priori presuppositions. For Cohen, "Kant discovered a new concept of experience," which consists in "the totality of synthetic propositions,

⁵ A11/B25.

⁶ KTE 279. Kant's concept of "transcendental" principles per se does not refer to the subject, according to Cohen's suggestion. Later, Fichte does refer "transcendental" to a subject, an absolute I.

⁷ Andrea Poma, *The Critical Philosophy of Hermann Cohen*, trans. John Denton (Albany: SUNY Press, 1997), 81.

⁸ This is in stark contrast with Schopenhauer's assessment of Kant.

⁹ I discuss this view in detail on pp. 00–00.

which form the content of mathematics and the pure natural sciences.”¹⁰ In the case of the critique of knowledge, the hard fact is provided by mathematics and the mathematical natural sciences. In other words, the assumed facts in mathematics and the natural sciences form the starting point of the transcendental inquiry. Our job as philosophers is to discover the underlying a priori laws that can adequately account for the objective validity of experience manifested in the sciences. Just as Kant assumes in the *Prolegomena* that the mathematical natural sciences provide a genuine knowledge of nature and goes on to suggest that there must be a set of a priori laws—synthetic a priori principles—that make this fact of science possible, Cohen is keen on identifying the necessary conditions of the experience that we accept as objectively valid as this expresses the fact of mathematical natural science.¹¹ A priori laws are “present” in this fact, and our job is to uncover these laws. The experience is given as a *task* (*aufgegeben*) to philosophy, so to speak.¹²

The transcendental cognition that is the goal of any transcendental inquiry must then be a priori. In Cohen there are three grades of the a priori, however.¹³ The first grade of the a priori is associated with the conceptual structure we can discover in our own mind by means of reflection. Space is a priori in the sense that it precedes all sensations and is at the basis of all outer experience. Space then is a priori in the sense of primary origin (*Ursprünglichkeit*), that is, it is conceptually prior to all other sensations. But this sense of “a priori” does not adequately account for the possibility of the apriority of space and time. This sense has nothing to do with a transcendental inquiry.

¹⁰ Jürgen Stolzenberg, “The Highest Principle and the Principle of Origin in Hermann Cohen’s Theoretical Philosophy,” in *Neo-Kantianism in Contemporary Philosophy*, ed. Rudolf A. Makreel and Sebastian Luft (Bloomington: Indiana University Press, 2009), 132–3.

¹¹ PIG 119–20.

¹² KTE 206.

¹³ For Cohen, the meaning of “pure” is related to the meaning of “a priori” but goes deeper. A thought is pure for Cohen when it is original and independent of experience. In this sense, pure thought would be the condition of the experience. Furthermore, pure thought is productive of the experience. As Cohen puts it, “pure thought in itself and only from itself must become the theory of knowledge” (LRE: W 13). Finally, pure thought is contentual: it has a content produced in and of itself. See Poma, *The Critical Philosophy of Hermann Cohen*, 80.

The second grade of the a priori is found on the plane of faculty psychology, that is, by way of the cognitive structures in the mind of the knowing subject. It consists in the forms of sensibility and the understanding.¹⁴ Space as a form of intuition is the same as an act of intuition,¹⁵ which Cohen sometimes calls "pure intuition." Space is thus a priori in the sense of the mode of intuition. But it is not a hypostasized organ—it is not physiological. Space then is not the psycho-physical organization of the subject.¹⁶ Nor is it innate (*angeboren*). It is not something inserted into our psyche at the time of birth. Thus Cohen definitely rejects the physiological program of H. v. Helmholtz and F. A. Lange as well as the psychological program advocated by J. F. Fries, J. F. Herbart and J. B. Meyer. In addition to space, Cohen also lists time and categories as exhibiting the second grade of the a priori. But these are not, in and of themselves, necessary for our purpose, namely, the transcendental account of the fact of science. In order to give an adequate account of experience, these a priori forms in the second grade are not necessarily required. The necessity of the second grade is derived from somewhere else. Cohen suggests that the second grade of the a priori is too psychological. Psychology in and of itself cannot be ultimately ground-laying (*letztbegründen*).¹⁷ Cohen also rejects Kant's transcendental deduction of the pure categories of the understanding as merely offering an empirical-psychological analysis of knowledge from subjective conditions.¹⁸

But these two grades of the a priori are grounded on the deepest level, that is, the third grade of the a priori, which consists in the "formal conditions of the possibility of experience." The concept of space is a priori in the third sense as well, that is, as a formal condition of the possibility of our experience. "That space is an a priori intuition,

¹⁴ Unlike Kant, Cohen does not separate sensibility and understanding. Sensibility and its product, intuition, cannot be separate from thought. Otherwise thought would fail to be productive. In this sense, Cohen follows J. G. Fichte and J. S. Beck.

¹⁵ KTE 46.

¹⁶ KTE 9.

¹⁷ Stolzenberg, "The Highest Principle," 133.

¹⁸ KTE 9.

after this clarification, now means: space is a constitutive condition of experience.”¹⁹ In other words, the third grade is identified with the laws that are constitutive of the possibility of experience. These a priori laws of human thought are what can ultimately explain the character of our experience of objects. Cohen then suggests that the principles of mathematics and the fundamental laws of pure natural science (mechanics) are a priori in this sense. As a consequence, these a priori laws “generate” objects of possible experience,²⁰ where experience is conceived as the “theories furnished by the mathematically precise nature of nature, considered as if laid out in printed books.”²¹ These are the laws of mathematics and mechanics that are considered independently of any particular knower. For Cohen, the Kantian language of cognitive activities must refer to the methods of mathematically precise natural science. Thus space and time are really the methods by which the mathematician constructs spatial magnitudes, and the categories really concern the method of the physicists constructing the representation of physical objects. In this way, Cohen gives a new foundation to the Kantian a priori.²² The a priori, whose possibility as a type of knowledge the transcendental inquiry must concern, does not simply precede the objects, but constructs them.²³

According to Cohen, the third grade of the a priori is relevant to Kant’s transcendental idealism: it shows the adequate sense in which space and time are not only empirically real but also at the same time transcendently ideal.²⁴ The third grade of the a priori does not consist in the cognitive structures in the subject’s mind in the physiological sense or

¹⁹ KTE 93.

²⁰ Klaus Köhnke, *The Rise of Neo-Kantianism* (New York: Cambridge University Press, 1991), 178–84.

²¹ KBE 27. See also Scott Edgar, “Hermann Cohen,” *The Stanford Encyclopedia of Philosophy* (Fall 2015 edition), ed. Edward N. Zalta, <http://plato.stanford.edu/archives/fall2015/entries/cohen/>.

²² KTE 13.

²³ KTE 48–9. See also Poma, *The Critical Philosophy of Hermann Cohen*, 13.

²⁴ Kant’s transcendental idealism here should be differentiated from empirical idealism, in both its dogmatic and problematic versions. Cohen then distinguished between two aspects of Kant’s transcendental idealism: concerning method, it is critical idealism, but concerning content, it is formal idealism (KTE 252).

in the psychological sense. Rather it consists in the logical sense—the principles of mathematics and the fundamental laws of pure natural science, that is, mechanics. As Poma suggests, “only in transcendental deduction can it find its full meaning, and its full justification as the ‘formal condition of the possibility of our experience.’”²⁵

Note that, for Cohen, the a priori in this strong sense is not discovered a priori. The starting point of Cohen's critical philosophy is the same as that of Kant: experience. As Kant puts it, “but although all our cognition commences *with* experience, yet it does not on that account all arise *from* experience.”²⁶ But experience here is not a mere “datum” in the empirical meaning of a prolix series of experiences.²⁷ Experience in this rich sense is the fact (*Faktum*) of the sciences.²⁸ Cohen claims that Kant discovered a new conception of experience.²⁹ The *Critique of Pure Reason* is basically a presentation of the distinctive Kantian theory of experience.³⁰ Kant's transcendental idealism is above all a new theory of experience. Critical philosophy is not concerned with the concept of knowledge in the common, generic sense nor the psychological sense of the cognitive process. It is rather concerned with pure knowledge, identified with the principles of science.³¹ As Cohen himself puts it, “we are going to start afresh. This means: we find ourselves again on the ground of the principles of the mathematical science of nature. They must again be indicated as pure knowledge and rediscovered in connection with logical reason.”³² The point of departure for any transcendental investigation then will always be experience. A proper transcendental investigation “starts from the fact of experience as a synthesis of

²⁵ Poma, *The Critical Philosophy of Hermann Cohen*, 15.

²⁶ B1.

²⁷ KTE 7.

²⁸ Sebastian Luft, “Reassessing Neo-Kantianism: Another Look at Hermann Cohen's Kant Interpretation,” *Dilthey: International Yearbook for Philosophy and the Human Sciences* 1 (2010): 5.

²⁹ KTE 3.

³⁰ KTE 5.

³¹ Poma, *The Critical Philosophy of Hermann Cohen*, 80.

³² LRE: W 11.

the phenomenon, only to return to the a priori conditions of the possibility of such a synthesis in sensibility and understanding, which then co-operate in the foundation of experience.”³³ But such a priori conditions naturally form a system of transcendental conditions whose highest principles serve as a genus for the rest of the conditions as the subordinate species.³⁴

Cohen on the Transcendental Method

Cohen suggests that his transcendental method adapts Kant’s analytic method to his purpose. Now, for Kant, an analytic method usually starts from what is reliably given to us as a fact and brings us to its relevant fundamental conditions by making a “backward” movement, that is, in a sort of “regressive” manner.³⁵ It thus moves from “what is more evident to us” toward “the supreme principle.”³⁶ In a typical case, it proceeds from the assumed body of knowledge to the necessary, enabling preconditions of this knowledge. Kant occasionally suggests that the analytic method would begin “from the sought as if it were given,” that is, from the existence of “a priori synthetic judgments” assumed as true, and also at the same time indicates that it would end with the only conditions under which such judgments are possible.³⁷ The conditions themselves are not established as independently true but only as “conditionally true,” depending on the truth of what is assumed. Accordingly, the method does not provide the conditioned with an independently authenticated ground by means of the now-discovered condition. It is thus truth-preserving—it merely preserves whatever truth the conditioned may originally contain. However, the kind of analysis involved in the analytic method is designed to abstract and isolate the underlying conditions—for

³³ LRE: W 12.

³⁴ KTE 138.

³⁵ See Halla Kim, *Kant and the Foundations of Morality* (Lanham, MD: Lexington Press, 2015), 7–9.

³⁶ G 4:392.

³⁷ P 4:277 n.

example, fundamental principles—from the body of our actual, ordinary (and well-entrenched) beliefs, as the only conditions under which the latter are possible. Thus the method may shed new light on the nature of the condition that was hidden from the mere consideration of the conditioned. The analytic method starts by assuming that “such cognitions from pure reason actually exist,” and in the case of theoretical cognition we typically appeal to two sciences, namely, pure mathematics and pure natural science. In this way, we proceed from such cognition to the “basis of its possibility.”³⁸

In the famous passages in the *Prolegomena* (e.g., §§4–5) Kant remarks that while the first Critique employs a synthetic (or progressive) method, the *Prolegomena* employs an analytic (or regressive) method.³⁹ Since the *Prolegomena* was written with an explicit intention to lay out the plan of the *Critique* in the form of a sketch by which the latter's central doctrines are restated and explained in simpler and more accessible terms, we may characterize the analytic method as that of an “exposition” as well. In other words, it helps us to gain deeper understanding of the conditioned (as well as the condition itself). The *Prolegomena* must therefore depend upon something already known to be trustworthy, from which we can set out with confidence and ascend to sources as yet unknown, the discovery of which will not only explain to us what we knew but also display the extent of many cognitions that all arise from the same sources. It starts with an assumed phenomenon, namely, the a priori synthetic nature of mathematics and physics, and ends with a conclusion about space and time as the only a priori forms of intuitions on the one hand and about the categories of the understanding on the other. Strictly speaking, the a priori synthetic status of mathematics and physics is not directly proved but rather presented and explicated conditionally. The obvious assumption on Kant's part here is that those doctrines have already received proper justification via what he calls “deduction” in the synthetically organized first Critique. Because the *Prolegomena* provides the necessary preliminary work for answering

³⁸ P 4:279.

³⁹ P 4:263, 274, 277 n.

the all too important question “How is metaphysics a priori possible?” it consequently paves the way for bringing the science of metaphysics, now conceived under this new light, into existence.⁴⁰

With this Kantian background, Cohen proceeds to employ the analytic method for his purpose. The task of critical philosophy is an investigation of the a priori elements of experience.⁴¹ Following Kant’s *Prolegomena*, Cohen suggests that the method of philosophical investigation is analytical. The analytic transcendental method thus must be at the heart of critical philosophy.⁴² Philosophical critique or rather its transcendental method holds facts in science and other cultural facts as the starting point of its investigation because they are deemed self-evident and thus valid.⁴³ It must be able to extract the a priori conditions of the necessity and strict universality of such facts. For Cohen, the *quid juris* question is the basic question for Kant’s philosophical enterprise in its entirety.⁴⁴ The legitimacy of our claims of knowledge must be subjected to transcendental investigation and can (potentially) be justified by a transcendental deduction. The different kinds of sciences (*Wissenschaften*) are “given” as objects for a “transcendental deduction,” as it were. The validity of our claims to knowledge must then be subjected to investigation by critical reason as well.⁴⁵ What Kant calls the principles of pure reason—the system of

⁴⁰ P 4:274.

⁴¹ Poma, *The Critical Philosophy of Hermann Cohen*, 8.

⁴² Michael Zank, “The Ethics in Hermann Cohen’s Philosophical System,” *Journal of Jewish Thought and Philosophy* 13, no. 1 (2004): 6. See also Poma, *The Critical Philosophy of Hermann Cohen*, 18.

⁴³ Zank, “The Ethics in Hermann Cohen’s Philosophical System,” 3.

⁴⁴ Ketil Bonaunet, *Hermann Cohen’s Kantian Philosophy of Religion* (Bern: Peter Lang, 2004), 10 n.; Beiser, *Genesis of Neo-Kantianism*, 466.

⁴⁵ Ethics and other facts of culture such as religion and art are no exception: they must be given as the phenomena of culture. They thus serve as natural objects for a transcendental deduction. The question of the justificatory ground (*Rechtsgrund*) for the existence and preservation of the facts of culture must be explored in a transcendental investigation. And facts of culture are not just a complex of theoretical judgments or beliefs, but also a comprehensive set of attitudes and ethical and ritual practices and institutions. Critical philosophy is to provide a rational justification in a priori transcendental cognition for our participation in such a complex of beliefs, attitudes, and practices. See Sebastian Luft, *The Space of Culture: Towards a Neo-Kantian Philosophy of Culture* (Cohen, Natorp, Cassirer) (New York: Oxford University Press, 2015), 28ff.; see also Bonaunet, *Hermann Cohen’s Kantian Philosophy of Religion*, 13.

synthetic a priori laws—are then such a priori conditions that presumably ground the fact of the sciences.

Cohen's analytic method, then, does not create "on its own, the science of the doctrine of nature,"⁴⁶ but must start out from the science as a "*factum*." The task of critical philosophy is not to construct a new rational cognition. Cognition in a given area is a fact, which is already out there, as "the fact of experience." Critical philosophy is transcendently-methodically committed to a *factum*. The analytic method must then lay bare the a priori conditions on which the validity of its pure cognitions rest. These conditions form the basis of mathematics and the pure natural sciences, and are embodied in the underlying synthetic a priori laws, which are called "principles of pure reason." However, this method provides no proof for the validity of mathematical-natural scientific cognition itself.⁴⁷ It can only extract those underlying conditions and describe them as lawfully valid.

The validity of the analytic method as the procedure for grounding our claims to knowledge had been challenged since Cohen's senior colleague Friedrich Albert Lange leveled a spirited charge of circularity against it: Cohen purports to seek that which makes possible experience or synthetic judgments necessary; but what he comes up with is simply the highest principles of experience.⁴⁸ These a priori conditions are claimed under the proviso that the same conditions are the explanatory ground for the presupposed experience of science.⁴⁹ The apparent reply on the part of Cohen is architectonic: he immediately suggests that these a priori conditions form the whole of the "transcendental system." The conditions then allegedly form the ground of mathematics and the natural sciences because they form the "total unified science." Each individual condition extracted by the analytic method is thus a part and parcel of a higher whole. This totality of transcendental conditions as unity hold together the unity of science.⁵⁰ For Cohen, such a highest

⁴⁶ KTE: W 577.

⁴⁷ Stolzenberg, "The Highest Principle," 138.

⁴⁸ KTE: W 138.

⁴⁹ KTE: W 135.

⁵⁰ KTE: W 135.

principle seems to be final and self-justified.⁵¹ For it presents the form of universal lawfulness. It thus appears that mathematical-natural scientific cognition is none other than what proceeds from such a principle.

However, in an important paper, J. Stolzenberg determines that this charge of circularity totally misunderstands Cohen's project. The charge is a red herring, so to speak, because the highest principle laid bare by the analytic method does not ground the experience. As Stolzenberg puts it, "because there is no *grounding* for the validity of experience, no vicious circle is set in motion with the exposition of a highest principle."⁵² Cohen must have never intended the analytic method as the full-fledged proof of the transcendental conditions.

For this reason, even though Cohen employs the Kantian analytic method, he cannot restrict himself to it. The Kantian analytic method that he adopts cannot provide any foundation for a tenable theory of knowledge. It cannot be ground-laying. It merely shifts the demand of justification further back. Thus Cohen cannot settle for the analytic method alone. His own transcendental method must improve on this Kantian notion of the analytic method and thus proceed from experience as the fact of science, that is, as natural scientific cognition, to its underlying presuppositions. The latter is pure rational cognition, which is typically presented in a system of principles. Yet Kant's elaborate theory of knowledge based on his transcendental inquiry involves merely an empirical-psychological analysis of the synthesis of knowledge into its subjective conditions.⁵³ The objectivity of knowledge will not be saved when it is founded on the subjective.⁵⁴ Kant's own attempt is inextricably psychological, because one cannot define

⁵¹ Stolzenberg, "The Highest Principle," 136.

⁵² Stolzenberg, "The Highest Principle," 139.

⁵³ Stolzenberg, "The Highest Principle," 133. Later in *The Bounds of Sense* (London: Methuen, 1996), P. F. Strawson engages in a spirited and sustained attack on the psychological tendency in Kant's program. In his view, Kant's theory of cognition deals with an "imaginary subject of transcendental psychology" (97). This criticism was exactly anticipated by Cohen's critique of Kant's transcendental project as too psychological. Cohen also rejects all psychological interpretations of Kant that were fashionable at the time.

⁵⁴ Stolzenberg, "The Highest Principle," 134.

the a priori as the origin of a concept from the sources of reason. Rather, the a priori elements must be solely built on the bare concept of the necessary and universal.

Kant and Cohen on the Synthetic Method

In an interesting work, Harry van der Linden points out that there is a decisive difference between Kant and Cohen in their conception of transcendental method. According to van der Linden, Cohen's transcendental method proceeds as follows: "[a] it searches for, and starts with, an X; [b] it looks for the rational presuppositions of the X; and [c] it argues that without these presuppositions, the X is unintelligible."⁵⁵ He then suggests that the "second step [b] is regressive and analytic in form."⁵⁶ This is clearly correct in view of our examination of the nature of the analytic method in the previous section. He then goes on to suggest that "the third step [c] is synthetic and provides a justification (deduction) for the right to use the presuppositions in question as well as all their logical consequences."⁵⁷ This is thoroughly misleading, however, as we will see shortly. There is no synthetic procedure or move involved in stating that the conditioned X would be unintelligible without these presuppositions. This is simply an integral part of the analytic method. What then is a synthetic method?

Kant famously attempts an analytic method in the *Prolegomena*, but in the *Critique of Pure Reason* itself the synthetic method is employed, because it treats nothing as given at the ground except reason itself and develops cognition out of its original spring without being supported by any *Faktum*. For Kant, then, a synthetic method moves "in the reverse direction" from the analytic method.⁵⁸ For example, the first Critique inquires into "pure reason itself and . . . determines in this source itself

⁵⁵ Harry van der Linden, *Kantian Ethics and Socialism* (Indianapolis: Hackett, 1988), 209.

⁵⁶ *Ibid.* 209.

⁵⁷ *Ibid.* 209.

⁵⁸ p 4:274–5.

the elements as well as the laws of its pure use according to principles.”⁵⁹ In a synthetic method, then, the facts themselves must be derived from concepts wholly *in abstracto*.⁶⁰ *The Critique of Pure Reason* searches for the sources of given sciences in reason itself.⁶¹ In this procedure, there is thus no fact or data given except reason itself and, in this respect, it seeks to “unfold cognition from its original germs without relying upon any facts.”⁶² As Kant succinctly puts in the *Lectures on Logic*, a synthetic method is one that proceeds “from principles to consequences” or “from the simple to the composite.”⁶³ The conditionally accepted truth of the claims given in the assumed body of knowledge receives its seal of warrant from the full-fledged proof of the a priori synthetic judgments. Here the conditional nature of the conditioned is finally discharged. This is because, in a synthetic method, the support for the assumed body of knowledge is elicited from “the source itself,” that is, “within pure reason itself,” and the “elements and laws of its pure employment” are determined accordingly.⁶⁴ Therefore, the synthetic method is truth-conferring.

While the analytic method seeks support in “something already known to be dependable,”⁶⁵ in the process of which the fundamental principle is “separated out” from the given experience without proving its independent “reality,” the synthetic method does the job of “truth-conferring” because it does not merely assume the truth of a priori synthetic cognition but instead makes an attempt to authenticate it “within the source itself,” that is, from the data of reason. Thus, even though both the analytical method and the synthetic method concern more than the inner, representational connections among our ideas, the former maintains the connection to reality conditionally, whereas the latter concerns the connections to reality in actuality, that is, it points to

⁵⁹ P 4:274.

⁶⁰ P 4:279. See Kim, *Kant and the Foundations of Morality*, 9–10.

⁶¹ P 4:280.

⁶² P 4:280.

⁶³ Kant, *Jäsche Logic in Lectures on Logic*, trans. Micheal Young (New York: Cambridge University Press, 1992), §117: AK 9:149.

⁶⁴ P 4:274.

⁶⁵ P 4:275.

the reality that those ideas have independently of their internal relations. The synthetic procedure is thus intrinsically ontological, as it involves a claim to the effect that, for a given topic T under discussion, there is really such a thing as T. In other words, it is laden with an ontological commitment in a way that an analytic procedure is not. For example, in Section III of the *Groundwork of the Metaphysics of Morals*, where a synthetic method is employed, we see that Kant attempts to show that the moral law is not chimerical but real, that is, actually valid for the human will if and only if it is free.⁶⁶

Likewise, in order to substantiate his claim about the grounding of being in thinking, Cohen needs a method other than the analytic one. Nevertheless, he cannot fall back on Kant's synthetic method, because for him a synthetic method is a psychological-empirical method of reconstructing the genesis of cognition from its subjective conditions, namely the faculty of reason, and should be rejected at all costs. It fails to achieve Kant's professed goal of "founding the necessity and strict universality of mathematical-natural scientific knowledge from a priori conditions."⁶⁷ It is now clear that Cohen is not only critical of any

⁶⁶This point is amply indicated by Kant already in the Preface to the *Groundwork*. There he remarks that the first two sections proceed "analytically from common knowledge to the determination of its supreme principle," but the third section proceeds "synthetically from the examination of this principle and its sources [in our practical reason] back to the common cognition" (G 4:392). In other words, the analytic method proceeds from the conditioned to the condition, but the synthetic method moves from the condition to the conditioned. Strictly speaking—and this is a point that can be easily misunderstood—the analytic method is not, from the viewpoint of human reason, independent of the synthetic method. The employment of the analytic method suggests that the principles obtained through the method will be subject to the synthetic method in due course. The nature of human reason is such that, when the principle, analytically abstracted, carries a conditioned validity, it seeks to discharge the condition by way of a synthetic method. Likewise, when we employ a synthetic method, we should first analytically operate on certain concepts or cognition, proceeding to their condition in a regressive manner, and then synthetically descend back to the conditioned from the condition. As a matter of fact, the *Critique of Pure Reason*, while proceeding synthetically from the outset, extensively employs an analytic method leading to an abstraction of some fundamental a priori synthetic propositions as the necessary conditions of our experience, and then methodically attempts to prove their independent validity in a synthetic manner. Thus, the analytic method and the synthetic method are not exclusive of each other but rather supplement each other when executed properly.

⁶⁷KBE: W 46; PIG: W 6.

psychological interpretation of Kant's transcendental philosophy but also of any psychological attempt to ground our claims to knowledge. The very science of psychology is not in a position to receive the rank of an ultimately ground-laying theory, exactly because it is an empirical science and thus devoid of any necessity and universal validity. At most, it can only provide hypothetically ultimate elements of consciousness. Philosophy understood as epistemology (or critique) aims at an ultimate grounding. Kant's synthetic method must be rejected as long as it is subjectivist, that is, insofar as it makes essential references to the modes of a subject. It thus appears that Cohen is just a few steps away from holding critical idealism without a subject.⁶⁸

The New Method of Transcendental Philosophy and the Concept of Origin

In order to understand the problematics facing Cohen here, we must remind ourselves that modern scientific knowledge presupposes basic mathematical principles that should be set in place before any investigation is implemented.⁶⁹ A proper employment of scientific method thus involves a number of elementary principles. These basic elements in turn govern the whole course of our scientific procedure and produce objective knowledge of nature in accordance with the method. For Cohen, philosophy, if properly performed, is to be a theory of science. In philosophy, a rich and fruitful philosophical procedure in pure cognition will require reflections on the underlying conditions that are concerned with those basic principles. All valid knowledge will then be regarded as the outcome of a systematic development of such logical elements. The

⁶⁸ I critically examine this view in the last section.

⁶⁹ Simon Fischer, *Revelatory Positivism: Barth's Earliest Theology and the Marburg School* (London: Oxford University Press, 1988), 40. We may call this Cohen's scientism, but in no way does it imply naturalism. Science determines the course of philosophical investigation for Cohen, but unlike Quine, Cohen never subscribes to naturalism. On the contrary, he is committed to strong metaphysical theses beyond the confines of science. For the anti-naturalist tendency in Cohen, see van der Linden, *Kantian Ethics and Socialism*, 210.

totality of the basic elements of pure cognition forms the transcendental ground of knowledge.

In a decisive step, Cohen postulates here that we stand in need of fundamental principles for our thinking to take off and proceeds to call these "hypotheses." For Cohen, a hypothesis in its Platonic employment means a thesis, which presents the point of logical commencement for a proper philosophical procedure. So the concept chiefly has a functional meaning. In this respect, any starting point for an argumentation with the purpose of grounding knowledge claims would be a hypothesis. A hypothesis, however, does not mean a valid proposition whose corroboration is somehow independently obtained. A hypothesis can be questioned whenever it is in conflict with other truths that we readily accept, and may be replaced by others. Now, Cohen proceeds to identify a hypothesis with what he terms "origin" (*Ursprung*). The principle of origin then must be ground-laying (*grundlegend*). As the ground-laying principle, origin must be presuppositionless, unconditioned and necessary: "a systematic point of culmination."⁷⁰ As a hypothesis, it is, however, subject to revision as well.⁷¹ This simply reflects the obvious historical fact that sciences make progress, and the old theories give way to the new.

Instead of a synthetic method, Cohen thus proposes a "method of hypothesis," which he develops in terms of the Platonic notion of hypothesis.⁷² In this respect, we may perhaps say that Cohen makes a transition from Kant's transcendental idealism to his own critical idealism by way of

⁷⁰ KTE: W 143.

⁷¹ We may then say that Cohen's overall transcendental method proceeds in two stages. The first, analytic stage, the deductive stage (Poma, *The Critical Philosophy of Hermann Cohen*, 94–6), identifies the ideal presuppositions in the natural sciences, e.g., the laws of the sciences. In the second stage, the productive stage (*ursprünglich*) of hypothesis, Cohen moves from the ideal presuppositions to the condition for their realization. The transition here is from the ideal to the real. Cf. Avi Bernstein-Nahar, *Accounting for Modern Jewish Identity: Hermann Cohen and the Ethics of Self-responsibility* (Ann Arbor, MI: University Microfilms, 1988). Cohen himself suggests that "the critique of cognition . . . separates sciences into presuppositions and foundations that are assumed in and for its laws."

⁷² Poma suggests "the dialectical method" as the name for this method (*The Critical Philosophy of Hermann Cohen*, 83–5), but in order to prevent it from being mistaken for a version of Hegelian dialectic, I think it is better to adopt "the method of hypothesis."

his recourse to Plato.⁷³ The Platonic concept of the idea as hypothesis means a thesis that forms the original point of argumentation.⁷⁴ Cohen thus suggests that “if the idea is primarily a hypothesis, then the category of origin is the most fundamental ground-laying [*Grundlegung*]; it is the foundation [*Grundlage*] of modern science.”⁷⁵ The idea that forms the basis of the mathematical and the pure natural sciences is only a hypothesis because its truth can in no way be guaranteed by the truths in such sciences.⁷⁶ Unless it is further proved (presumably from the deliverances from our reason), its grounding capacity alone will not confer any final truth on it. When we produce cognition of nature, our cognition is grounded on our thinking. On this conception, all beings are beings that are posited in thinking. As the ground of being, thinking is ground-laying (*Grundlegung*) of being.⁷⁷ Cohen immediately identifies this ground-laying thinking with origin, and goes on to suggest that the principle of origin must be able to present the laying of the ground of being through thinking only by means of its form of universal lawfulness.⁷⁸

From this perspective, it is no wonder that “the methodological center is the idea of hypothesis, which we have developed into the judgment and the logic of origin.”⁷⁹ At the center of this system are the highest principles as summed in the principle of origin. This origin is the foundation of modern science, because as the form of universal lawfulness it is an unconditioned condition of our knowledge. The principle of origin then lays the ground for the fact of the sciences. Thus, it is due

⁷³ Cohen’s ethics in ERW therefore is the strict application of Kant’s method to the establishment of the moral law in practical philosophy. See Halla Kim, “Hermann Cohen and the Foundations of Ethics,” in *Proceedings of the 12th Kant Congress* (Berlin: de Gruyter, forthcoming); and “Hermann Cohen on the Concept of Law in Ethics,” in *Jewish Religious and Philosophical Ethics*, ed. C. Hutt, Halla Kim, and B. D. Lerner (London: Routledge, forthcoming).

⁷⁴ Stolzenberg, “The Highest Principle,” 140.

⁷⁵ LRE: W 597.

⁷⁶ Stolzenberg, “The Highest Principle,” 134.

⁷⁷ LRE: W 145.

⁷⁸ LRE: W 145.

⁷⁹ LRE: W 601.

to origin that the identity of thinking and being is made possible. Indeed, this principle is what makes being what it is.

If, however, we emphasize the Platonic heritage in the notion of idea as hypothesis, then we will end with the autocracy of being. For Cohen, being is not autonomous: it cannot dictate our thought. It has meaning and value only as being posited and created in thinking. Being is not given to us passively but actively assigned by thinking. Cohen sometimes speaks of the "eternity of reason" in elaborating on the idea of determination of being by means of thinking. Cohen's new method of hypothesis thus is also based on the Kantian insight that we can know a priori of objects only what we have put into them.⁸⁰ We can see that in his attempt to validate knowledge of nature, Cohen emphasizes the transcendental activity of thinking: "thinking itself is the goal and object [*Gegenstand*] of its activity."⁸¹ "Only thinking itself can generate what validity counts as being."⁸²

For Cohen, then, thinking determines reality. Knowledge and being originate in thinking. This naturally leads to his view that cognition depends on the judgment of origin (*Das Urteil des Ursprungs*). Even though this notion is cryptic, there is no denying that this serves as the foundation of his system. We may say that Cohen's refusal to go with any appeal to the synthetic method could perhaps be explained by his appropriation of the conception of origin. What then is origin? We can characterize its essential features in five points:

(1) An origin is a point at the beginning of a line of argumentation where cognition takes off.⁸³ As a beginning of thought, it is the first judgment.⁸⁴ Furthermore, it is the principle of every judgment. As Cohen

⁸⁰ Bxii, Bxiii; KTE 112. This is sometimes called the "Vico formula" (Luft, *The Space of Culture*, 46).

⁸¹ LRE: W 29.

⁸² LRE: W 81.

⁸³ In this characterization I follow Fischer, *Revelatory Positivism*, 39–42.

⁸⁴ Cohen proceeds to offer four types of judgments:

1. The judgments of the law of thought (*Denkgesetz*), of which there are origin, identity, contradiction.
2. The mathematical judgments, of which there are reality, plurality, totality.

puts it, “origin is not only the necessary beginning of thought; it must act as the moving principle in every development.”⁸⁵ As such, the principle of origin is the foundation of all pure knowledge. As logic of thinking, the logic of origin is “in itself logic of pure knowledge” (*an sich selbst die Logik der reinen Erkenntnis*).⁸⁶

(2) Origin expresses the incorrigible ideal of the form of universal lawfulness immanent in all claims to knowledge. In this sense, origin is an unconditioned condition for our knowledge and being. It is the principle that lays the ground for all our knowledge. As Cohen puts it, “if knowledge is the same as principle, it is conditioned by origin.” And if the “thought is the thought of knowledge, then its beginning and ground are in the thought of origin.”⁸⁷

(3) Origin is potency. It is an active principle that produces knowledge and being from its own inner resources. Human thought is creative and productive of the objects on its own: “without origin, a principle (a logical principle or some mathematical hypothesis of natural science) cannot be productive.”⁸⁸ His theory of origin thus shows the unrestricted sovereignty of thinking, where thinking is the potency to procure the prerogatives of the determination. Origin is thus unconditioned as this potency. It qualifies the conditioned as that which is. Thinking characterizes itself as what it is and it qualifies that which it objectifies. Thinking thus has its character as principle and serves the function of

3. The judgments of the mathematical natural sciences, of which there are substance, law, concept.

4. The methodical judgments, of which there are possibility, actuality, necessity.

This corresponds to Kant’s table of categories, but for Cohen judgments take precedence over categories. For Kant, each category is identified by its own type of judgment, but this is not necessarily the case in Cohen as several categories may determine one and the same judgment (LRE: W 50). Judgment invariably expresses the unity of thought but the categories may change, reflecting the progress of the sciences. See Poma, *The Critical Philosophy of Hermann Cohen*, 88.

⁸⁵ LRE: W 36.

⁸⁶ LRE: W 38.

⁸⁷ LRE: W 36.

⁸⁸ LRE: W 227.

determination: it determines something determinable as something. Determination then means a formation of judgment.

(4) Thinking of origin performs the function of a signpost, indicating the path generative thinking follows in attaining knowledge and being.⁸⁹ In this sense, an origin represents a transition or rather development in our construction of any claims to knowledge. An origin is a law for thought (*Denkgesetz*) and in this regard it stipulates the trajectory for pure thinking to traverse. The origin then determines the law-governed (*gesetzlich*) process of generative thinking, just as a mathematical equation determines a curve. The principle of origin is thus the principle of continuity (of determination). Continuity is a law of thinking (*Die Kontinuität ist ein Denkgesetz*).⁹⁰ In this process of continuity, determination and determinability are connected to each other. Cohen calls the function of judgment which founds the establishment of continuity “the principle of continuity.” Continuity then indicates the generative law of thinking in knowledge. We can see that thinking of origin combines with generative thinking to produce knowledge and being. Cognition is pure when it is obtained by the determination of thinking and by the determination which functions as predication. Predication establishes unity.⁹¹ Unity of knowledge is also unity of determination. But the latter is not only the formal connection but it also has objective meaning (*sachliche Bedeutung*). For thinking generates being as an object. This point then leads to our final point.

(5) Origin is an activity. As Cohen puts it, pure thinking is the “thinking of origin” (*Denken des Ursprung*).⁹² Origin is not a ready-made product, nor is it a stationary state of an entity. Rather it works as it actively operates qua the principle of grounding for all our knowledge. It is productive, and not merely representative. It thus grounds being in and of itself. In this way, the identity of thinking and being—or rather

⁸⁹ Fischer, *Revelatory Positivism*, 41.

⁹⁰ LRE: W 91.

⁹¹ LRE: W 47.

⁹² LRE: W 36.

the grounding of being in thinking—is achieved. Cohen thus suggests that “foundations [*Grundlagen*] are ground-layings [*Grundlegungen*].”⁹³ Further, defying the Kantian distinction between form (i.e., activity) and matter (i.e., content), Cohen suggests that “activity itself is content.”⁹⁴ What thinking produces is its own content, which does not depart from itself.⁹⁵ The content of thought is not given as the manifold but must be actively produced by thinking. Thus we may say: “the activity of production itself is the product.”

This last point has an important repercussion for our understanding of Cohen’s view of the generation of objects in thinking. Kant once suggested that, while the form of knowledge comes from the mind, its matter must be given from without, beyond the confines of the mind. Cohen cannot accept this duality of the workmanship of knowledge: “there is no other way to discover the object than that offered by the unity of knowledge. It represents the unity of the object. And unity of knowledge is formed in the unity of judgment. That is how we acquire this determination of judgment: the unity of judgment is the formation of the unity of the object in the unity of knowledge.”⁹⁶ The object for Cohen is then produced by thinking as the determinable. The first demand of thinking is that it places the “origin of each and every content that it is able to create into thinking itself.”⁹⁷ In a dramatic passage, Cohen claims that “thinking has to discover being in origin” (*daß das Denken im Ursprung das Sein zu entdecken hat*).⁹⁸ Being then owes its meaning to the work of the determination in judgment. Cohen thus has this to say about this point: “nothing can count as given for pure thought. It must also produce the given by itself.”⁹⁹ An object then is

⁹³ ERW: W 84.

⁹⁴ LRE: W 60.

⁹⁵ LRE: W 29.

⁹⁶ LRE: W 60.

⁹⁷ LRE: W 82.

⁹⁸ LRE: W 36.

⁹⁹ LRE: W 101.

nothing more than a determinable something. And the determination of something determinable proceeds by way of limitation.¹⁰⁰ In the first place, the determinable something cannot be anything but limitedly determined. The initial constitutive function of determination or judgment is “nothingness” (*Nichts*).¹⁰¹ Determination is applied to the determinable only via “nothingness,” via an opposition to each and every determined. It thus indicates an infinite relation. It is a limitation.¹⁰² Therefore, judgment of origin is infinite or limitative.

But we should note that Cohen's view on the generation of the object exhibits some subtle development over his long career. In his early period, in a manner reminiscent of Jakob Sigismund Beck, Cohen presents the Analytic as the central component of the Critique. The Aesthetic is modeled on the Analytic. The important consequence of this view is that the object in space is not simply given to us before the categories are applied to it.¹⁰³ On the contrary, the object is given exactly because the subject synthesizes it in accordance with the categories. We can receive the manifold of sensation only when the categories have done their proper work. Cohen gives a limited role to the given. Thus he says: “things are appearances, but are they phantasms? Not at all! Appearances are, inasmuch as there are laws, in which the reality of appearances is grounded and in which the reality of appearances subsists.”¹⁰⁴ Thought realizes the given. The given is thought-laden. At this point Cohen suggests that “law is reality [*die Realität*], which means reality is to be conceived of as an abstract thought.”¹⁰⁵ Then he continues: “the so called things have their reality in the aggregate of the laws of appearances; they (the laws) are appearances . . . Appearances are truly objects [*Objekte*]. They alone are the

¹⁰⁰ For Spinoza, “*determinatio negatio est*” (determination is negation) (*Epistle*, 50 in *Opera*, IV). Hegel concurs as well.

¹⁰¹ LRE: W 49.

¹⁰² LRE: W 58.

¹⁰³ KTE 179–81.

¹⁰⁴ KBE 20.

¹⁰⁵ KBE 21.

real things which, though the laws of pure thought, become definite objects [*Gegenstände*] of intuition.”¹⁰⁶

In his later period, however, in the *Logik der reinen Erkenntnis*, Cohen presents a more radical view according to which the object is nothing but a generated nexus of laws.¹⁰⁷ Sensation does not play any role in the generation of knowledge and being. Laws are the sole distinguishing feature of known objects.

This doctrine of origin in Cohen suggests that the principle of origin guides our transcendental activities exclusively by means of judgments. The theory of origin turns out to be a theory of judgment (*Urteilsdenken*).¹⁰⁸ Judgment is the fundamental form of thinking or a fundamental form of determination (*der Grundform des Denkens, der Grundform der Determination*). Judgment as a function of determination is the specific structure of unity. The synthesis of unity exhibits the moments of separation, unification, and preservation. However, the synthesis performed by judgment does not proceed by way of a combination of material as in Kant. It is not a composition of a given manifold.¹⁰⁹ Instead the synthesis is of unity, as unity of separation (*Sonderung*) and unification (*Vereinigung*).¹¹⁰ It is through judgment “according to which and in itself [*nach und in welchem*] the discovery of the object takes place.”¹¹¹ Judgment as the productive process of thought then proceeds by way of separation and unification.¹¹²

It cannot be denied that for Cohen origin has the status of hypothesis and can only signify a laying of grounds, which in principle is capable of revision.¹¹³ This simply suggests that the system of pure

¹⁰⁶ KBE 23.

¹⁰⁷ Fischer, *Revelatory Positivism*, 44.

¹⁰⁸ Werner Flach, “Cohen’s *Ursprungsdenken*,” in *Hermann Cohen’s Critical Idealism*, ed. Reiner Munk (Leiden: Springer, 2005), 45.

¹⁰⁹ LRE: W 26.

¹¹⁰ Flach, “Cohen’s *Ursprungsdenken*,” 46.

¹¹¹ LRE: W 47–8.

¹¹² LRE: W 61.

¹¹³ Geert Edel, “Kantianismus oder Platonismus? Hypothesis als Grundbegriff der Philosophie Cohens,” *Il Cannoichiale: Rivista di studi filosofici*, 1–2 (1991): 59–87.

cognition that is the foundation of our knowledge of the world is not fixed once and for all but must be “conceived as a system that is incomplete, open, and capable not only of expansion but also of revision.”¹¹⁴ Indeed, Cohen adamantly and forcibly points out that “completeness” with respect to the “number of categories” would not be a fullness, but would form “an open wound of logic.”¹¹⁵ Modern mathematical-natural science presents the *Factum* of becoming (*Werdefaktum*).¹¹⁶ The inventory of natural-scientific cognitions is not complete but in the process of continuous development, expansion, or revision. The progress of science implies the progress of pure cognition.¹¹⁷

Now, even though the principles as a hypothesis may be revised, the very relation of being to thought is not subject to revision. The principle of origin as the matrix for the idea of the form of lawfulness is not revisable.¹¹⁸ The idea of the form of lawfulness is at the basis of pure cognition. But the principle of origin is the principle of any given law, that is, the principle of original positing of being attributable to thinking.¹¹⁹ The idea of lawfulness is not revisable. Cohen notes: “the ultimate foundations of logic are ground-layings whose expressions must change according to the progress of the problems and of the insights . . . the eternity of reason is confirmed in the historical nexus of ground-layings.”¹²⁰ The origin then must be an inexorable idea of ground-laying, no matter how its expressions may change with respect to the progress of science. As Cohen puts it, “the basic form of being is the basic form of judgment,” and the basic form of judgment is the “basic form of thinking.”¹²¹ Our judgment has its ground of validity solely in

¹¹⁴ KTE: W 519. See Stolzenberg, “The Highest Principle,” 141.

¹¹⁵ LRE: W 396.

¹¹⁶ LRE: W 76.

¹¹⁷ LRE: W 396.

¹¹⁸ Stolzenberg, “The Highest Principle,” 138–9, emphasizes this point.

¹¹⁹ LRE: W 144.

¹²⁰ LRE: W 245.

¹²¹ LRE: W 47.

thinking. Therefore thinking and being are the same. Being is created in and through thinking. Such thinking is “thinking of the origin” (*Denken des Ursprungs*).¹²² Thinking of origin is the sole origin of being. This is not changeable nor revisable.

In sum, then, we may say that the thought of principles serves a foundation for our claims to knowledge because it is grounded in origin. This is what Cohen calls the logic of origin in the *Logik der reinen Erkenntnis*.¹²³ The principle of origin thus works as the ground-laying principle of knowledge. Cohen’s theory of origin is part of his doctrine of the principles of validation of our knowledge. This is why his doctrine of principles is at the same time a doctrine of principles of thinking and a doctrine of the objective foundations of the sciences.

Critical Idealism Without a Subject?

We have seen that Cohen, like Kant before him, never calls the validity of scientific knowledge into question. Cohen never treats the challenge of global skepticism seriously and always conceives the latter (in its moment of opposition to dogmatism) as a mere stepping stone on the way to critical idealism.¹²⁴ Cohen in effect asks how the cognition of the world is possible, given its inescapable manifestation in the world. His “critique” is concerned mainly with the necessary presuppositions of scientific knowledge.

Cohen, however, seems to reject Kant’s transcendental subject together with its ineluctable features of human psychology. For Kant, a law can organize the material supplied by empirical intuition to produce objective knowledge of appearances solely within the bounds of the unity of consciousness. This unity is provided by transcendental apperception.¹²⁵ The Kantian transcendental subject thus enables the laws to combine with sensations to generate objective knowledge.

¹²² LRE: W 36.

¹²³ LRE: W 31–2.

¹²⁴ Poma, *The Critical Philosophy of Hermann Cohen*, 60.

¹²⁵ A107.

Unconditionally giving up on a synthetic method because of its psychologistic implications, Cohen starts with the fact of the sciences (*Faktum der Wissenschaften*), the positive scientific knowledge of the world, as the inescapable truths of the natural and human sciences. Transcendental deduction then would proceed by way of reflection upon experience, in this robust sense, without any recourse to mental faculties.¹²⁶ In this way, Cohen can avoid commitment to any dubious metaphysical claims concerning the human subject. He simply substitutes the fact of the sciences for the fact of subjective experience. As opposed to Kant, who starts with the experience of the ordinary human subject and ends with claims about the cognizing subject's faculty psychology (sensibilities and understanding), Cohen starts from the actual sciences and ends with their necessary presuppositions. If Kant sought the conditions for the possibility of experience in the ordinary sense, Cohen then seeks the conditions for the validity of the sciences.¹²⁷ Cohen's deduction thus remains strictly within the bounds of scientific knowledge and consists exclusively in the development of a system of judgments and their logical presuppositions (*Urteilen und Grundbegriffen*) as the foundation for the actual sciences (*Wissenschaftsarten*).¹²⁸

It thus appears that there is no room for the subject and its workings in Cohen's new epistemology. In his view, critical philosophy must conduct an inquiry into the possibility of justification of human knowledge on the plane of the novel concept of experience, that is, of experience as the best of the sciences currently affords it. It thus investigates the conditions of the possibility of the exact natural sciences.¹²⁹ Only in this way can the philosophical critique achieve a sustainable justification of claims to a priori cognition. It therefore appears that Cohen has effectively ditched the idea of the subject.

¹²⁶ Bernstein-Nahar, *Accounting for Modern Jewish Identity*, 89; Geert Edel, *Von der Vernunftkritik zur Erkenntnislogik* (Freiburg: Karl Alber, 1988), 64.

¹²⁷ Geert Edel, "Cohen und die analytische Philosophie der Gegenwart," in *Philosophisches Denken—Politisches Wirken: Hermann-Cohen-Kolloquium Marburg 1992*, ed. Reinhard Brandt and Franz Orlik (Hildesheim: Georg Olms, 1993), 197.

¹²⁸ Edel, "Kantianismus oder Platonismus?" 64.

¹²⁹ Luft, "Reassessing Neo-Kantianism," 4.

As Brelage puts it, “critical idealism is not idealism of the subject, but of idea; it owes its name, not to Locke’s idea in its meaning as representation, but to Plato’s Idea, in its interpretation as hypothesis, law, a principle of validity. This law—and not the spirit—is the absolute in which all knowledge is grounded. It is the subject of all knowledge, which confers objectivity on the latter.”¹³⁰ His philosophy thus would be critical idealism without a subject.

This interpretation, however, must be resisted. Cohen’s view of the subject is clearly not psychological-subjective, but this does not mean that he dispenses with the idea of the subject altogether. This interpretation does not see that Cohen’s conception of the subject is intimately connected to the justification of knowledge. The subject in Cohen is operating as the source of principles that are required to validate knowledge. His method, namely, the transcendental method, has been designed to justify the experience enacted in scientific experiments. The method then must reconstruct the logical steps involved in the construction of scientific cognition, which necessarily involves the activities of the operating subject.¹³¹ For Cohen’s critical idealism, knowledge is not representation but judgment. It is not a passive mirror of nature but an active enterprise. Thus Cohen’s transcendental subject must be distinguished from the object of psychology. It is what actively exercises valid judgments. This subject then must be differentiated from the concrete psychological subject. Instead it operates as the ideal signpost which guides the trajectory of human knowledge. As Poma puts it, “Cohen’s transcendental subject, his knowing or scientific consciousness, is therefore distinct from the concrete psychological subject; rather it is the ideal, the norm, and the task for the latter, i.e., the concrete empirical self.”¹³² As such, it forms a system of principles on which knowledge is rationally justified. Cohen thus does not abandon the notion of the subject altogether; he instead

¹³⁰ Attributed to Nicolai Hartmann. See also Manfred Brelage, “Transzendentalphilosophie und konkrete Subjektivität,” in *Studien zu Transzendental Philosophie* (Berlin: de Gruyter, 1965), 97. For a different view, see Poma, *Critical Philosophy*, 62.

¹³¹ Fischer, *Revelatory Positivism*, 28.

¹³² Poma, *The Critical Philosophy of Hermann Cohen*, 63.

advocates a new conception of it. In this way, knowledge is not a representation of beings in themselves, unilaterally dictated to us by the order of being. This system of principles is thus not only independent of the things in themselves but also independent of any individual subject.¹³³ This is how Cohen can subscribe to Kant's Copernican Revolution without committing himself to the metaphysical duality of the Kantian type of transcendental idealism. We may then say that at the center of Cohen's system is a transcendental (cultural) subject. But this subject is not an individual human agent. Nor does it deduce the empirical I's.¹³⁴ But we can certainly deduce from this subject the idea of the sciences in accordance with which the practical in the relevant realm is governed. Cohen's system is a transcendental logic of self-determination, at the center of which lies a distinctive transcendental subject.

The presence of the subject is more dramatically manifested in his practical philosophy. For Cohen, human cognition is not only concerned with natural sciences (*Naturwissenschaft*) but also jurisprudence (*Rechtswissenschaft*). Just as Cohen's logic (or theory of knowledge: *Erkenntnistheorie*) is the area in which the transcendental method is applied to the *Faktum* of natural sciences (*Naturwissenschaft*), ethics is the area in which the transcendental method is applied to the *Faktum* of jurisprudence. In ethics, the ideal condition or presupposition that the transcendental method is concerned with is identified with the ideal self (*das Selbst, das Ich, der Person, der Mensch*). While in Cohen's logic the idea is posited as hypothesis, in practical philosophy, such a hypothesis is replaced by an ideal as an end that is to be fulfilled infinitely by the ideal self and which can thus impose a limit on our ways of being.¹³⁵

¹³³ Poma, *The Critical Philosophy of Hermann Cohen*, 64.

¹³⁴ Bernstein-Nahar, *Accounting for Modern Jewish Identity*, 85.

¹³⁵ In this respect, Cohen's idealism is a critical but not absolute idealism, and it is thus differentiated from Fichte's idealism. Fichte rids the subject of its psychological associations, but his I still remains metaphysical without being fully transcendental. As Cohen puts it, "for Kant the transcendental means the criterion of the a priori; while for Fichte it means self-consciousness in all its psychological ambiguity" (KBE 290). In this respect, Fichte loses contact with Newton's science of nature. Fichte's idealism is thus a subjectivism of a self-indulging I, which goes back to Descartes's innatism (Poma, *The Critical Philosophy of Hermann Cohen*, 74). Further, Fichte does not make a distinction between logic and ethics, between being (*Sein*) and what ought to be (*Sollen*).

The origin of ethics in Cohen then turns out to be the idea of humanity (*Menschheit*), the ideal self.¹³⁶

Conclusion

How transcendental is Cohen's critical idealism? First, his critical idealism is transcendental because it employs the analytic method. Indeed, Cohen derives the necessary a priori conditions for the objectivity of mathematical-natural knowledge to his satisfaction in what he calls the principle of origin. Yet he does not stop there and proceeds to remove all psychological implications from this investigation. The a priori conditions in question have nothing to do with any particular cognitive subject, and they turn out to be the laws of mathematics and mechanics considered independently of any individual knower. Furthermore, the structure of the analytic method does not aim to justify these necessary a priori conditions in and of themselves. This is because the objective truth of the fact of the sciences does not preclude its underlying premise from being false, nor does it preclude the multiplicity of mutually incompatible premises.¹³⁷ The starting point of transcendental inquiry must be "self-evident,"¹³⁸ but the conditions themselves are not self-evident. This forces Cohen to admit that the highest principle of the transcendental cognitions must be originary at a deeper level. For this move to be fruitful, Cohen must introduce the method of hypothesis and the ensuing concept of the principle of origin. But strictly speaking this move goes beyond the narrow conception of transcendental method—that is, the analytic method—that Cohen finds in Kant. In this respect, Cohen's transcendental inquiry goes well beyond its Kantian counterpart.¹³⁹

The noumenon ceases to be an idea or task and becomes a real principle. In Fichte, logic is grounded in ethics, being in what ought to be. Cohen cannot accept this despotism of the subject.

¹³⁶ van der Linden, *Kantian Ethics and Socialism*, 207.

¹³⁷ Stolzenberg, "The Highest Principle," 134.

¹³⁸ Zank, "The Ethics in Hermann Cohen's Philosophical System," 8.

¹³⁹ I would like to thank Steven Hoeltzel and Sebastian Luft for comments on this chapter.

Halla Kim is Professor of Philosophy and a faculty member at the Schwalb Center for Israel and Jewish Studies at the University of Nebraska at Omaha. Recent publications include his book *Kant and the Foundations of Morality* (Lexington Books, 2015), as well as an anthology (with S. Hoeltzel), *Kant, Fichte, and the Legacy of Transcendental Idealism* (Lexington Books, 2014). Presently he is also editing *Jewish Religious and Philosophical Ethics*, together with C. Hutt and B. D. Lerner (Routledge, forthcoming). Among his research interests are Kant, German idealism, modern Jewish thought, and East Asian philosophy.

8

Heidegger's Failure to Overcome Transcendental Philosophy

Eric S. Nelson

The Problem of Transcendental Philosophy

John Searle has complained that “it ought to arouse our suspicions that people who spend enormous efforts on interpreting [Heidegger’s] work disagree on the fundamental question whether he was an idealist.”¹ Scholars of Heidegger’s philosophy have similarly been unable to agree whether or to what extent he was committed to transcendental philosophy, which Kant defined as the analysis of the necessary conditions of possible experience in

¹ John R. Searle (ed.), “The Phenomenological Illusion,” in *Philosophy in a New Century: Selected Essays* (Cambridge: Cambridge University Press, 2008), 107.

E.S. Nelson (✉)

Department of Philosophy, Hong Kong University of Science and Technology,
Clear Water Bay, Hong Kong

general, or whether he overcame it in a “radical” new thinking of being and its history.²

The continuing disagreement concerning the role of transcendental philosophy in his thought can be attributed in large part to Heidegger himself. It reflects his shifting and ultimately inconsistent positions concerning his relationship with the transcendental heritage. In this chapter, I trace Heidegger’s changing and ambiguous relationship with “transcendental philosophy,” which he defined at various points in relation to the philosophy of the subject and subjectivity, reflective-representational thinking, and the horizontal understanding of meaning.

As he recounted in the 1963 lecture “My Way into Phenomenology,” Heidegger’s philosophical training was deeply shaped by the transcendental philosophies of Neo-Kantianism (Heinrich Rickert) and phenomenology (Edmund Husserl) that he studied at the University of Freiburg.³ Despite or perhaps because of this education, Heidegger would repeatedly endeavor to distance himself from and break with the transcendental paradigm of his Freiburg teachers, Rickert and Husserl, while still tacitly relying upon it and at times—even in his later works when it should have long been overcome—reverting to its language and argumentative strategies.⁴

²For instance, Hubert L. Dreyfus stresses Heidegger’s break with transcendental thought: “Heidegger developed his hermeneutic phenomenology in opposition to Husserl’s transcendental phenomenology.” See *Being-in-the-world: A Commentary on Heidegger’s Being and Time, Division I* (Cambridge: MIT Press, 1991), 2. Cristina Lafont emphasizes the anti-transcendental nature of Heidegger’s “linguistic idealism” and the ontological difference, which makes the distinction between the empirical and transcendental impossible, in *Heidegger, Language, and World-Disclosure* (Cambridge: Cambridge University Press, 2000), 17. Skepticism about Heidegger’s success at overcoming transcendental philosophy is developed by Karl-Otto Apel; see “Meaning-Constitution and Justification of Validity: Has Heidegger Overcome Transcendental Philosophy by History of Being?” in Karl-Otto Apel (ed.) *From a Transcendental-Semiotic Point of View* (Manchester: Manchester University Press, 1998), 103–21. A cogent case for Heidegger’s continuity with transcendental philosophy is made in Daniel Dahlstrom, “Heidegger’s Transcendentalism,” *Research in Phenomenology* 35, no. 1 (2005): 29–54. These are two Heideggerian voices, according to Steven G. Crowell, *Husserl, Heidegger, and the Space of Meaning: Paths toward Transcendental Phenomenology* (Evanston: Northwestern University Press, 2001), 9.

³GA 14:93–101.

⁴On Heidegger’s relation to Husserl, see Leslie MacAvoy, “Heidegger and Husserl,” in *Bloomsbury Companion to Heidegger*, rev. edn, ed. François Raffoul and Eric S. Nelson (London: Bloomsbury, 2016), 135–42.

Heidegger engaged in a number of attempts to reformulate transcendental philosophy, such as in terms of fundamental ontology and world-disclosure in the second half of the 1920s, and to break with transcendental philosophy. An early attempt to disentangle himself from the transcendental paradigm can be seen in his early post-war turn toward existence- and life-philosophy and hermeneutics, which he developed in particular through his reading of Dilthey.⁵ Heidegger attempted in his “hermeneutics of factual life” to overcome transcendental philosophy and what he depicted as its static, ahistorical conception of the constitution of meaning, through an interpretive-existential analysis of concrete situated existence.⁶ His lecture courses of the early 1920s promised a radical breakthrough and return to life in its very facticity. Theodore Kisiel has described how Heidegger's early project was adjusted through his reappropriation of transcendental philosophy as fundamentally ontological and his explicit return to transcendental philosophy, ontologically understood, in the mid-1920s during the period of *Being and Time* (1927) and *Kant and the Problem of Metaphysics* (1929).⁷ During

⁵ For a detailed discussion of Dilthey's significance for the early Heidegger, see Eric S. Nelson, “The World Picture and its Conflict in Dilthey and Heidegger,” *Humana Mente: Journal of Philosophical Studies* 18 (2011): 19–38; Eric S. Nelson, “Heidegger and Dilthey: Language, History, and Hermeneutics,” in *Horizons of Authenticity in Phenomenology, Existentialism, and Moral Psychology*, ed. Hans Pedersen and Megan Altman (Dordrecht: Springer, 2015), 109–28.

⁶ On the context of Heidegger's hermeneutics of facticity, see Theodore Kisiel, “On the Genesis of Heidegger's Formally Indicative Hermeneutics of Facticity,” in *Rethinking Facticity*, ed. François Raffoul and Eric S. Nelson (Albany: State University of New York Press, 2008), 41–67; Eric S. Nelson, “Questioning Practice: Heidegger, Historicity, and the Hermeneutics of Facticity,” *Philosophy Today* 44 (2001): 150–9.

⁷ Cf. Theodore Kisiel, *The Genesis of Heidegger's Being and Time* (Berkeley: University of California Press, 1993), 9. Kisiel depicts the development of Heidegger's *Being and Time* through “three drafts”: (i) the “hermeneutical” or “Dilthey”-influenced draft that reflects his early project of a hermeneutics of factual life (1915–21); (ii) a “phenomenological-ontological” draft that relies on working through Aristotle's ontology and a renewed engagement and struggle with Husserl's phenomenology (1921–24); and (iii) a quasi-transcendental and Kantian draft (1924–27). *Being and Time's* failure motivated Heidegger's movement away from Kant toward a renewed thinking of the anti-transcendental ontological motivations, such as the “it worlds” and the primordial happening (“*es ereignet sich*”) and upsurge of a pre-intentional and pre-theoretical “it” (*es*) or “there” (*da*), of the first draft without its existential and life-philosophical dimensions.

this period, Heidegger identified the inner truth of transcendental philosophy with the fundamental ontology of being.

A second example of Heidegger's break with transcendental philosophy, and the most frequently disputed example in the literature, is the so-called "turning" (*Kebr*) in the mid-1930s, understood as an attempt to overcome the lingering transcendental character of *Being and Time*, which is concerned with the "ontology of Dasein" or, "in Kantian terms," a "preliminary ontological analytic of the subjectivity of the subject."⁸ According to Heidegger's later self-interpretation in the 1930s, *Being and Time* had failed to address—or was in being-historical "errancy" (*seinsgeschichtliche Irre*) concerning—the genuine question of being (*Seinsfrage*) by overemphasizing the constitutive role of the subject and its distinctive temporality. The thereness of "being-there" (*Dasein*) was not yet thought radically enough. Heidegger identified transcendental philosophy after the turn, linking it with his wider "history of being" (*Geschichte des Seins*), with the priority of the subject and subjectivity that he associated with problems of modernity—rooted in the origins and historical unfolding of Western metaphysics—in works such as "The Age of the World Picture" (1938).⁹

After the "turn," Heidegger continued to alternate between the rhetoric of radically overcoming transcendental philosophy—for its subjective, horizontal, reflective-representational, and modernist character—and the possibility of an alternative conception of transcendental constitution that occurred through the "event" (*Ereignis*) of being, world, and history, rather than through the analysis of the conditions of possibility of the subject. Heidegger did not overcome his ambiguous relationship with the transcendental tradition and could not overcome transcendental philosophy. This clarifies why, despite his own self-interpretations, the transcendental interpretation of Heidegger's context and—both early and later—works remains trenchant.¹⁰

⁸ SZ 24. Page numbers for SZ refer to the German edition.

⁹ Martin Heidegger, *Off the Beaten Track* (Cambridge: Cambridge University Press, 2002), 83–4. On the problematic of the subject and subjectivity in Heidegger, see François Raffoul, *Heidegger and the Subject* (Atlantic Highlands, NJ: Humanities Press, 1998).

¹⁰ For transcendental approaches to Heidegger's thought, see for instance: *Transcendental Heidegger*, ed. Steven G. Crowell and Jeff Malpas (Stanford: Stanford University Press, 2007); Daniel Dahlstrom,

Phenomenology as Idealism and Transcendental Philosophy

Searle's suspicions concerning the implicit idealism of Heidegger's phenomenology have significant historical precedents. In one of the earliest critiques of the phenomenology articulated in *Being and Time*, Georg Misch in his *Lebensphilosophie und Phänomenologie: Eine Auseinandersetzung der Diltheyschen Richtung mit Heidegger und Husserl* argued for the inherently subjective-idealistic character of the phenomenological movement. Misch, who was familiar with the development of Heidegger's thought throughout this period, identified Heidegger's position with a subjective, Fichtean-style ethical "idealism of freedom." Günther Anders would rechristen it an "idealism of unfreedom" (*Idealismus der Unfreiheit*) in his 1937 critique of Heidegger's philosophy in relation to his involvement with National Socialism.¹¹ Misch's description corresponds to Heidegger's "metaphysics of freedom" phase during the late 1920s and early 1930s, which was unfolded through his interpretation of German idealism and Friedrich Wilhelm Joseph Schelling in particular.¹² Indeed, instead of keeping his distance, Heidegger had an affirmative sense of the achievements of German idealism, as a philosophical elevation from which later generations have fallen.¹³

"Heidegger's Transcendentalism," *Research in Phenomenology* 35, no. 1 (2005): 29–54; Dermot Moran, "Dasein as Transcendence in Heidegger and the Critique of Husserl," in *Heidegger in the Twenty-First Century*, ed. Tziovani Georgakis and Paul J. Ennis (Dordrecht: Springer, 2015), 23–45.

¹¹ Misch described how Heidegger "is ethical-idealistically positioned, while an objective idealistic orientation is revealed in Dilthey (*ethisch-idealistisch eingestellt ist, während bei Dilthey die objektiv-idealistische Einstellung sich darin verriete*)," in Georg Misch, *Lebensphilosophie und Phänomenologie: Eine Auseinandersetzung der Diltheyschen Richtung mit Heidegger und Husserl*, 2nd edn (Leipzig: Teubner, 1931), 29–30. Cf. Günther Anders, *Über Heidegger* (München: Beck, 2001), 28; see also the discussion of Misch's critique of Heidegger's idealism in Eric S. Nelson, "Dilthey, Heidegger und die Hermeneutik des faktischen Lebens," in *Dilthey's Werk und seine Wirkung*, ed. Gunter Scholtz (Göttingen: Vandenhoeck & Ruprecht, 2013), 103.

¹² On Schelling's significance for Heidegger's thinking of freedom and imagination in this key period of transition between *Being and Time* and the turn, see Christopher S. Yates, *The Poetic Imagination in Heidegger and Schelling* (London: Bloomsbury, 2015).

¹³ See GA 40:34; Martin Heidegger, *Introduction to Metaphysics* (New Haven: Yale University Press, 2014), 45.

Jürgen Habermas, whose 1954 doctoral work was on the contradictions between history and the absolute in Schelling, connected Heidegger's thought with idealism, which unlike classical German idealism relativizes rather than grounds rational knowledge. Habermas repeatedly depicted Heidegger's early phenomenology of Dasein as a subjective decisionism and his later thinking of being as a form of "linguistic idealism" that prioritizes "the world-disclosing function of language."¹⁴ Habermas concluded that Heidegger is beholden to the worst elements of the idealistic heritage, temporalizing and relativizing it, and is unable to take an inter-subjective and communicative turn that would rehabilitate the rational claims of transcendental philosophy.¹⁵

Searle maintains in his essay "The Phenomenological Illusion" a position concerning phenomenology not unlike that of Habermas in *The Philosophical Discourse of Modernity*. He argues that in actual fact the entire classical phenomenological tradition is committed to semantic idealism or, if we abandon the claim of idealism as overly polemical, a perspectival reduction of knowledge and truth claims to a point of view or language game. Searle describes the semantic position thus: "a view is idealist in this semantic sense if it does not allow for irreducibly *de re* references to objects. All references to objects are interpreted as being within the scope of some phenomenological operator."¹⁶ According to Searle's argument, this description encompasses not only Husserl's conception of the fundamental character of the intentionality of consciousness but also less obviously idealistic operators such as Dasein (being-there) in Heidegger or the body in Merleau-Ponty.

Searle's argument appears at first glance overly simplistic and in need of complication, given the notions of passivity and

¹⁴ Jürgen Habermas, *The Postnational Constellation: Political Essays* (Cambridge: Polity Press, 2001), 146. Also see Jürgen Habermas, *Der philosophische Diskurs der Moderne: Zwölf Vorlesungen* (Frankfurt: Suhrkamp, 1985), 168.

¹⁵ William D. Blattner has extensively argued that Heidegger is a "temporal idealist." See his *Heidegger's Temporal Idealism* (Cambridge: Cambridge University Press, 1999).

¹⁶ Searle, "Phenomenological Illusion," 107. For his depiction of phenomenology as a semantically idealist or quasi-idealistic perspectivalism, see 128–32.

sedimentation in Husserl; of facticity, thrownness, and being beyond Dasein's operations of meaning and sense-making in Heidegger; of the entwinement of the body in the flesh of the world in Merleau-Ponty; or of the priority of alterity in Levinas. Classical phenomenology cannot be identified with subjective or objective idealism, even if it might well be explicitly (as in Husserl) or implicitly committed to underlying premises of transcendental philosophy.¹⁷

Searle asserts that Husserl's transcendental philosophy emphasizes meaning ("what is said") over reality ("the thing" itself): "all of his talk about the transcendental ego and the primacy of consciousness is . . . a part of his rejection of the idea that what I have been calling the basic facts are really basic."¹⁸ Searle claims in addition: "transcendental subjectivity for Husserl does not depend on the basic facts; rather, it is the other way round."¹⁹ Searle's description conflates a necessary condition of *x* with the reduction of *x* to that necessary condition; that is, the idea that the intentional and proto-intentional constitution of meaning is necessary for there to be meaningful facts for us (what he calls the semantic) with the idea that real things are predetermined and constructed through the constitution of meaning (what he calls the *de re* independent reality).

The notions of transcendental conditions in Kant and transcendental constitution in Husserl did not entail the rejection of experientially encountering things and scientifically explaining the empirical world, which correspond to the pre-predicative experience of the life-world (*Lebenswelt*), on the one hand, and the theoretical idealization of the sciences, on the other hand, as shown in Husserl's late works such as *Experience and Judgment* and *The Crisis*

¹⁷ I argue that Heidegger and Levinas never overcome the premises of transcendental philosophy in Eric S. Nelson, "Biological and Historical Life: Heidegger between Levinas and Dilthey," in *The Science, Politics, and Ontology of Life-philosophy*, ed. Scott M. Campbell and Paul W. Bruno (London: Continuum, 2013), 15–29.

¹⁸ Searle, "Phenomenological Illusion," 124.

¹⁹ Searle, "Phenomenological Illusion," 125.

of *European Sciences and Transcendental Phenomenology* (1936). Husserl demonstrated how this is possible without appealing to the metaphysical realist's mystical positing of a nonconstituted or unexperienced and uninterpreted reality (the *de re* disconnected from the *de dicto*).

To complicate the picture further, Searle's criticisms of Husserl overlap with Heidegger's and those of other later phenomenological critics: the fundamental intentionality and relationality of consciousness, without which there would not be meaning, acting or knowing for Husserl, is construed as predetermining the entirety of reality when it is making the experience and interpretation of reality possible. Intentionality does not isolate the ego or mind in itself; it designates the irreducible constitutive relationality that allows humans to encounter things and "know facts" in meaningful ways precisely because consciousness has relational and intentional characteristics that make these processes possible.

Phenomenology cannot be semantic idealism in Searle's sense. Husserl did of course describe his phenomenology as transcendental idealism and transcendental subjectivism. This indicates the difference between: (i) idealism₁ as the constitution of all—including material and natural—reality out of the subject, or the ideational (semantic) nature of all reality (which Husserl never maintained); and (ii) idealism₂ as the constitution of sense and meaning from the fundamental non-dual co-relationality of subject and object disclosed through the phenomenological reductions:

Consciousness describes how the world becomes manifest: The attempt to conceive the universe of true being as something lying outside the universe of possible consciousness, possible knowledge, possible evidence, the two being related to one another merely externally by a rigid law, is nonsensical. They belong together essentially; and as belonging together essentially, they are also concretely one, one in the only absolute concretion; transcendental subjectivity.²⁰

²⁰ Edmund Husserl, *Cartesian Meditations: An Introduction to Phenomenology* (Dordrecht: Springer, 1977), 84.

Heidegger's Ambiguous Relationship with Transcendental Philosophy

Heidegger repeatedly sought to redefine and/or break from and overcome the idealism of his mentors and of the Western metaphysical tradition. In these attempts, it is evident that Heidegger—who declared himself a follower of the “realist” Husserl of the *Logical Investigations* (1900–01)—shared some of Searle's concerns about the idealizing tendencies of Husserl's phenomenology throughout his lecture courses of the 1920s and his later break with the transcendentalism of *Being and Time* itself.

Being and Time, however Heidegger's thought is ultimately interpreted, is clearly composed in the context of the project of transcendental philosophy, as can be seen in the text itself. His account of the temporality of Dasein in *Being and Time* aimed at elucidating “time as the transcendental horizon for the question of being.”²¹ The language of “transcendental horizons” is borrowed from Husserl. The term “horizon” indicates how the transcendental dimension cannot consist purely in a description of the activities of the ego and subject, to the extent that the subject is referred to ever wider conditions and horizons of meaning-constitution and genesis. In Heidegger's argument, time cannot be said to be constituted by the subject, as existentially reinterpreted temporality constitutes the very sensibility of Dasein's being from the unchosen thrownness of birth to the inappropriable facticity of death, which indicates two occasions that define human Dasein while defying the meaning and sense-making activities of the subject.

Nonetheless, in *Being and Time* and *Kant and the Problem of Metaphysics*, the notion of the transcendental is pushed in other directions that Husserl rejected as a betrayal of transcendental phenomenology for the sake of a renewed metaphysics (in his critique of Heidegger's ontological language) and philosophical anthropology (in his criticism of Heidegger's use of existential language).²² Husserl argued that *Being*

²¹ SZ 39.

²² See Edmund Husserl, *Psychological and Transcendental Phenomenology and the Confrontation with Heidegger (1927–1931)*, Collected Works, vol. 6 (The Hague: Nijhoff, 1997).

and Time failed because of its departure from, rather than its lingering commitment to, the philosophy of transcendental subjectivity. Heidegger had reified phenomenology into an existential anthropology.²³

The continuing affinities and the growing distances between Husserl and Heidegger are apparent in the latter's interpretation of the transcendental in passages such as this one: "*Being is the transcendens pure and simple*. The transcendence of the being of Dasein is a distinctive one since in it lies the possibility and necessity of the most radical *individuation*. Every disclosure of being as the *transcendens* is *transcendental* knowledge. Phenomenological truth (disclosedness of being) is *veritas transcendentalis*."²⁴ What is noteworthy about this passage in this context is that Heidegger is harkening to the scholastic, presumably more ontological, sense of "transcendental" in this passage and connecting it with the more or less existential question of the singular and unique being of the self in its individuation. Heidegger's 1915 qualifying dissertation had concerned the problem of the relation between universal categories of meaning and the form of individuation (the *haecceitas* or "thisness") in scholasticism.²⁵ As "existential," categories are only meaningful insofar as they are enacted and embodied in diverse ways of being; as "ontological," they concern the question of how being is and how beings are rather than the issue of how to access epistemically their reality through knowledge.

Neo-Kantian and Husserlian transcendental philosophy had for Heidegger prioritized the question of knowledge and the knowing subject in its concern for the logical and epistemic conditions of possibility.²⁶ It is

²³ Husserl, *Psychological and Transcendental Phenomenology*, 505; see also Edmund Husserl, "Phänomenologie und Anthropologie," in *Aufsätze und Vorträge (1922–1937)*, *Gesammelte Werke XXVII*, ed. Thomas Nenon and Hans Rainer Sepp (Dordrecht: Kluwer, 1989), 164–81. See my discussion in Eric S. Nelson, "What Is Missing? The Incompleteness and Failure of Heidegger's Being and Time," in *Division III of Heidegger's Being and Time: The Unanswered Question of Being*, ed. Lee Braver (Cambridge: MIT Press, 2015), 210.

²⁴ SZ 38.

²⁵ GA 1:203, 253. On the significance of *haecceitas* for Heidegger, see John van Buren, *The Young Heidegger: Rumor of the Hidden King* (Bloomington: Indiana University Press, 1994), 105–7.

²⁶ See, for instance, GA 58:180.

this understanding of the complicity of the transcendental, subjectivity, and the epistemological essence of modern philosophy since Descartes that Heidegger would place into question by rethinking them in more primordial ontological and life-existential senses in the 1920s and in endeavoring to confront and overcome them altogether in his later thinking.

Heidegger's controversial redescription of Kant's conception of the transcendental as concerned with the ontological question of being instead of the epistemological problem of possible knowledge developed through his readings of medieval scholastic philosophy, its antecedent origins in Aristotle, and his perhaps polemical reinterpretation of Kant's *Critique of Pure Reason* as an ontological work, which brought him into dispute with the Neo-Kantian reading of Kant. Heidegger would debate with Ernst Cassirer over the very character of the critical philosophy in the Swiss town of Davos in 1929.²⁷ Kant's transcendental philosophy becomes in Heidegger's reading a general ontology: nothing less than the ontological determination of the region of all beings.²⁸ The question concerning possible conditions is primarily a question of being (*Seinsfrage*) that Kant had failed to pose radically enough and with adequate self-understanding.

Heidegger rethought the question of the conditions of possibility as the question of the possibility of essence in his confrontation with Kant's critical philosophy, which would soon turn toward German idealism as a step beyond Kant. In the context of his reading of Hegel and Schelling, transcendental philosophy will no longer be rethought as ontology. Heidegger would abandon the project of *Being and Time* (1927) and *Kant and the Problem of Metaphysics* (1929) as insufficiently radical. Transcendental philosophy belongs to the very forgetting of the ontological question of

²⁷ On the historical context and intellectual implications of the Davos debates, see Michael Friedman, *A Parting of the Ways: Carnap, Cassirer, and Heidegger* (LaSalle: Open Court Publishing, 2000); Peter Eli Gordon, *Continental Divide: Heidegger, Cassirer, Davos* (Cambridge: Harvard University Press, 2010).

²⁸ GA 25:58. Compare Frank Schalow, "Heidegger and Kant: Three Guiding Questions," in *Bloomsbury Companion to Heidegger*, rev. edn, ed. François Raffoul and Eric S. Nelson (London: Bloomsbury, 2016), 105–11. Compare my discussion of Schalow's analysis of Heidegger's Kant in Nelson, "What Is Missing?" 202.

being, as it is tied to the abstract understanding of subjectivity introduced through René Descartes's conception of the ego *cogito*. Subjectivity, and the transcendental philosophy associated with it as the "philosophy of subjectivity," are inherently bound together and to be overcome, for the mature Heidegger.²⁹

To the extent that *Being and Time* is still indebted to and haunted by the transcendental perspective, it is a failure and necessarily incomplete, and yet, in Heidegger's self-interpretation, it is a meaningful step towards his thinking and history of being that emerged in the 1930s.³⁰ Even though *Being and Time* attempted to overcome an ahistorical and abstract consciousness in favor of historical-concrete existence, it remained within the reifying and false abstractions of the tendency of transcendental thinking.³¹ Transcendental philosophy is only one more name among others for the philosophy of the modern subject that is the culmination of the history of metaphysics and its forgetfulness of being.³²

Heidegger and advocates of his later thought maintained that his new transformative thinking signifies an overcoming of transcendental philosophy.³³ Each time Heidegger further developed his thinking, he

²⁹ On transcendental philosophy as the philosophy of subjectivity, see GA 14:96; on subjectivity and modernity, see GA 67:242.

³⁰ On Heidegger's conception of the history of being, see Eric S. Nelson, "History as Decision and Event in Heidegger," *Arbe* 4, no. 8 (2007): 97–115; Eric S. Nelson, "Heidegger, Levinas, and the Other of History," in *Between Levinas and Heidegger*, ed. John Drabinski and Eric S. Nelson (Albany: State University of New York Press, 2014), 51–72.

³¹ Martin Heidegger, *Briefe an Max Müller und andere Dokumente*, ed. Holger Zaborowski and Anton Bösl (Freiburg and München: Alber, 2003), 102.

³² GA 48:75. For a justification of idealism as a crucial interpretation of modern self-reflexivity and freedom, see Robert B. Pippin, *Idealism as Modernism: Hegelian Variations* (Cambridge: Cambridge University Press, 1997).

³³ On Husserl's critique of Heidegger, see Steven G. Crowell, "Does the Husserl/Heidegger Feud Rest on a Mistake? An Essay on Psychological and Transcendental Phenomenology," *Husserl Studies* 18, no. 2 (2002): 123–40; Sebastian Luft, "Husserl's Concept of the 'transcendental person': Another Look at the Husserl–Heidegger Relationship," *International Journal of Philosophical Studies* 13, no. 2 (2005): 141–77. On the transition from *Dasein* to *Sein* and Heidegger's later critique of the philosophy of the subject, compare Bret W. Davis, *Heidegger and the Will: On the Way to Gelassenheit* (Evanston: Northwestern University Press, 2006), 197; Patricia J. Huntington, *Ecstatic Subjects, Utopia, and Recognition: Kristeva, Heidegger, Irigaray* (Albany: State University of New York Press, 1998), 188.

associated his previous phases with the lingering traces of transcendental philosophy. It should be asked: Did Heidegger overcome transcendental philosophy in and through the turn and his mature thought, and is it necessary to overcome it? A critical reading of the entirety of his works indicates that it is question-worthy whether he ever overcame his ambiguous relationship with his transcendental heritage. It is perhaps the case that transcendental philosophy continues to be the best way to make sense of his projects despite his own anti-transcendental self-interpretations.

Three Attempts to Rethink Transcendental Philosophy

We will now briefly and schematically consider three examples of Heidegger's attempts to reconsider transcendental thought: his early project of a hermeneutics of facticity, his thinking of world-constitution and disclosure as an approach to the question of meaning, and his being-historical thinking of the event and history of being, which is supposed to overcome his earlier transcendentalism as well as that of the western metaphysical tradition.

The Hermeneutics of Factual Life

First, beginning with his early post-war interest in life-philosophy and hermeneutics (Dilthey), Heidegger attempted to "overcome" transcendental philosophy and its "static," ahistorical, idealizing conception of constitution through a hermeneutics of factual life, only to return to an explicitly transcendental-horizonal language in the period of *Being and Time* and *Kant and the Problem of Metaphysics*. It is noteworthy that he recurrently rejected transcendental philosophy and then would return to it by employing its language and strategies. An early example of this ambivalence is his early project of a hermeneutics of facticity or factual life. He contrasted a philosophy that sought to clarify the factual conditions of life and lived-experience with the dead and empty

abstractions of a transcendental philosophy that was intrinsically limited in its capacity to articulate the structures of factual life immanently from out of itself.³⁴

For instance, he claimed in an early criticism of Husserl that “the being is pure because it is defined as *ideal*; that is, *not real* being.”³⁵ Husserl’s reductions draw from “the initially given concrete individuation of a stream of experience what is called the pure field of consciousness, that is, a field which is no longer concrete and individual, but pure.”³⁶ Thus, in emphasizing ideality instead of facticity, transcendental phenomenology cannot encounter and draw from the real being and concreteness of the entity in question.³⁷

Heidegger’s early argumentation employs the lived and interpretive “categories of life”—a Diltheyan idea that contests the ahistorical, static categories of consciousness and the reductive interpretation of reason offered by traditional transcendental philosophy, and a forerunner to the existential categories or “existentials” unfolded in *Being and Time*—and, more fundamentally, *logos*; that is, the communicative event and enactment of factual existence in and through the constitutive medium of language.³⁸ It is in this situation that Heidegger redefined the transcendental dimension of his project by the mid-1920s as an existential-ontological one, which would provide a prior basis to history and nature, interpretation and explanation, by exploring the quasi-transcendental constitution of human existence as communication, event, and enactment.³⁹

The “concrete” issues of life are not addressed by remaining at that level of understanding of consciousness, for Heidegger. They demand a radically ontological thinking achieved through strategies such as “formal

³⁴ For instance, see GA 60:13.

³⁵ GA 20:145–6; translation from Martin Heidegger, *History of the Concept of Time: Prolegomena* (Bloomington: Indiana University Press, 1985), 106.

³⁶ GA 20:138–9; translation from *History of the Concept of Time*, 101.

³⁷ GA 20:146–8; see *History of the Concept of Time*, 106–8.

³⁸ See my “Biological and Historical Life,” 22.

³⁹ GA 61:173.

indication” and hermeneutical anticipation that poses the question of one’s own life/existence in relation to the question of the meaning of being. This would be articulated through his hermeneutics of authentic and inauthentic speech, logos as revealing and as concealing, which he unfolded in confrontation with Aristotle’s *Rhetoric* in the lecture courses on the *Basic Concepts of Aristotelian Philosophy* (1924) and *Plato’s Sophist* (1925).⁴⁰

Heidegger’s thinking of the originariness of language in the mid-1920s distinguished his project from other prevalent forms of transcendental philosophy that focused on the constitutive role of consciousness and the self. This indicates the fundamental role of his thinking of language and, as we will see next, of world as well, which Searle has depicted as perspectival semanticism. It is, however, the residual transcendental character of Heidegger’s thinking of language, world, and being that releases it from Searle’s critique.

Transcendental Philosophy and World

The relationship between world-constitution and transcendental philosophy offers a second example of Heidegger’s shifting attitude toward the transcendental paradigm. In the late 1920s, he identified the truth of transcendental philosophy with ontology and time as the condition of the finite subject. Time is the transcendental horizon of the question of being posed by Dasein. Is this a temporal idealism, as Blattner has argued?⁴¹ It is, to the extent that it is Dasein’s originary temporality that clarifies the notion of time that happens to it, even if it happens as inappropriate in moments such as birth and death.

In *The Metaphysical Foundations of Logic*, a lecture course from 1928, Heidegger analyzes the transcendental character of world in Kant’s first Critique as ontologically what cannot be part of another world or be

⁴⁰ A particularly helpful work for exploring the priority of language in the 1920s is Scott M. Campbell, *The Early Heidegger’s Philosophy of Life: Facticity, Being, and Language* (New York: Fordham University Press, 2012).

⁴¹ Blattner, *Heidegger’s Temporal Idealism*.

reduced to any possible ontic physical world.⁴² This is an account of the world as a meaningful whole that is irreducible to any part and a transcendental thesis about the world that remains the point of departure for Heidegger's later thinking of world.⁴³

Heidegger is seen in this same lecture course refining his depiction of the transcendental paradigm by introducing his conceptions of the transcendental dispersion and dissemination of neutral Dasein into non-neutral ways of being through thrownness.⁴⁴ Thrownness is not the ontic fact of being born but rather the transcendental-ontological structure disclosed through the facticity of ontic birth. Human finitude and mortality, which remain crucial to Heidegger's later understanding of human existence between heaven and earth, are not merely anthropological facts of human life; they are a transcendental-ontological condition of meaningfulness and meaninglessness, instead of a bare, finite, empirical duration of time.

In *Kant and the Problem of Metaphysics*, transcendental philosophy is inquiry into the understanding of being emerging from Dasein's ecstatic transcendence. In this context, especially in the late 1920s in works such as "On the Essence of Ground" (1928/29), the transcendental signifies Dasein's transcending, or surpassing and overstepping, of limiting structures and the exposure of being-in-a-world: that is, "transcendence as being-in-the-world" and "world co-constitutes the unitary structure of transcendence."⁴⁵

In "On the Essence of Ground," Heidegger stated that "world co-constitutes the unitary structure of transcendence; as belonging to this structure, the concept of world may be called transcendental. This term names all that belongs essentially to transcendence and bears its intrinsic possibility."⁴⁶ The co-relational character of mortal and immortal, world and earth, in their difference and conflict does not eliminate the sense of

⁴² See the discussion in GA 26:224–5; translation in Martin Heidegger, *The Metaphysical Foundations of Logic* (Bloomington: Indiana University Press, 1984), 175.

⁴³ For an extended assessment of the problematic of world in Heidegger, see Lafont, *Heidegger, Language, and World-Disclosure*.

⁴⁴ See GA 26:173–4; translation in *Metaphysical Foundations of Logic*, 138.

⁴⁵ Translation from Martin Heidegger, *Pathmarks* (Cambridge: Cambridge University Press, 1998), 109.

⁴⁶ Heidegger, *Pathmarks*, 109.

the whole that Heidegger here described as transcendental. The idea of transcendental method continues to be heard in his later notion of way (*Weg*) as world-disclosing, which happens to a wanderer on the way to whom the world is being disclosed. Heidegger's poetic thinking resonates with the poetic dimensions of the transcendental tradition at work in Hölderlin and Schelling. World is a meaningful whole: "world is not a mere collection of the things—countable and uncountable, known and unknown—that are present at hand," Heidegger noted in "The Origin of the Work of Art," but world is fundamentally world opening and the worlding of the world.⁴⁷ The openness of world is a reinterpretation of the phenomenological account of world-constitution. The world remains constitutive of meaning, even in the modern epoch in which the world is darkened and the meaningfulness of things seemingly lost.

Heidegger's Thinking of Being

We should now turn to the history and event of being and consider whether it has overcome the transcendental paradigm. The turn in the mid-1930s was interpreted by Heidegger and his subsequent supporters as a radical break with the lingering transcendental character of *Being and Time*, which continued to forget the question of being by over-emphasizing the role of the subject in the form of "being-there" (*Dasein*). Transcendental philosophy became increasingly identified after the turn, linking it with the history of being, with the priority of the subject and subjectivity that were associated with the problems of modernity. As I discuss in another place, the post-turn Heidegger persistently conjoins and critiques the transcendental conception of the subjectivity of the subject and the modern experience of the priority of the subject as an ahistorical and worldless reification.⁴⁸

⁴⁷ Translation from Heidegger, *Off the Beaten Track*, 23.

⁴⁸ Nelson, "What Is Missing?" 211. On the transcendental sense of subjectivity, and posing the question of subjectivity more radically, compare Heidegger, GA 2:24, 106, 229, 382; GA 26:129, 160, 190, 205, 211; GA 27:11. For Heidegger's later critique of the subject, see GA 5:243; GA 69:44; GA 79:101, 139.

Even after 1935 and the turn toward being, Heidegger continues to hesitate at times between the rhetoric of radically overcoming transcendental philosophy and the possibility of an alternative, desubjectified conception of transcendental constitution occurring through the impersonal constitutive media of history, language, and world, rather than constitution occurring through the activity of the subject or Dasein. In a passage from 1936/37, later published in *Elucidations of Hölderlin's Poetry*, Heidegger describes the significance of the constitutive interplay of history, language, and world for human life: "only where there is language, is there world, that is, the constantly changing cycle of decision and work, of action and responsibility, but also of arbitrariness and turmoil, decay and confusion. Only where world holds sway is there history."⁴⁹

History, language, and world are not accidental attributes found in empirical life to be explained by scientific inquiry or understood through lived-experience. Heidegger sharply distinguishes between contingent ontic-empirical histories (*Historie*) and an ontologically (that is, in the last analysis despite Heidegger's intentions, transcendently) conditioning history (*Geschichte*) of being. Language has a transcendental character in conditioning, structuring, and speaking through the speaker. The world happens to the finite mortal human subject who is called to wait and listen for the murmurs of being.

Heidegger's later thinking might be interpreted as a philosophy of the transcendental passivity of the subject. Still, it is Dasein that exists in and from the between (*Zwischen*) of the ontological difference between being and entities. This openness and its play of concealment and unconcealment, in which being is a transcendental field of uncertainty, cannot be fully disclosed or understood and can even be crossed out and erased. This thinking transforms and yet cannot fully displace *Being and Time's* transcendental horizon and conditions of the question concerning the sense of being.

According to Heidegger, innerworldly beings are only disclosed and manifest in the context of the world-clearing of being in which humans find themselves and which happens prior to any transcendental thinking or

⁴⁹ GA 4:38; translation from Martin Heidegger, *Elucidations of Hölderlin's Poetry* (Amherst, NY: Humanity Books, 2000), 56.

projection by Dasein or the human subject.⁵⁰ The transcendental is no longer located in the conditions and structures of subjectivity, much less of an epistemological character of a knowing subject. Nonetheless, history's historicizing, language's speaking, and the world's worlding have a transcendental character in contrast with, for instance, accounts that explain them as ontic phenomena through natural and material constitution.

Heidegger's mature thinking of world-opening and world-clearing as an ontological event would be more primordial than and prior to—and yet, however, resonates with—the idea of transcendental constitution understood as the conditions of possibility for the relationship between world and being, on the one hand, and the finite mortal human subject, on the other.

Conclusion

Heidegger's problematic and tense relationship with transcendental philosophy shapes his early efforts to transform and overcome it, as well as the trajectory of the question concerning the meaning of being. Based on the indications briefly discussed above, Heidegger did not embrace or turn to the historically available alternatives to transcendental philosophy, such as naturalism and materialism (e.g., physical constitution) or historicism and social constructivism (e.g., social-cultural constitution). The model of the transcendental conditions and constitution of the meaningfulness of the world reverberates throughout his later thinking of a more originary and primordial world-disclosure and world-clearing.

Meaningfulness appears to be no longer explicitly tied to the meaning-making activity of the human subject, as being becomes the necessary condition of possible meaningfulness for humans.⁵¹ Being always surpasses

⁵⁰ Compare Martin Heidegger, *The Piety of Thinking: Essays* (Indianapolis: Indiana University Press, 1976), 95.

⁵¹ For an account of Heidegger that stresses the priority of being over any act of meaning or sense, and the irreducibility of *Sein* to *Sinn*, see Richard Capobianco, *Heidegger's Way of Being* (Toronto: University of Toronto Press, 2014). On the priority of *Sinn* in interpreting *Sein* in Heidegger, see Thomas Sheehan, *Making Sense of Heidegger: A Paradigm Shift* (Lanham, MD: Rowman and Littlefield, 2015).

its historical epochs and human perspectives, such that it is irreducible to semantic perspectivalism or idealism. Yet it is mortal humans who interpret, poeticize, and dwell in response to and in the context of this event of being. The non-dual and irreducible relationality of subject and object as a condition of meaning links Heidegger's later thinking with transcendental phenomenology at the same time as he disorients the Husserlian subject.

It is accordingly questionable whether Heidegger: (i) eliminated the constitutive role of meaning and the subject in projecting it onto being and the history of its event and epochs, which he persistently distinguished from ontic-empirical histories; (ii) achieved a coherent and experientially appropriate nonrepresentational and non-horizonal notion of the sense and meaning of being; and (iii) articulated a philosophy that could be "naturalized" without undermining its very structure and sense that prioritizes the event of meaning.⁵²

As argued in this chapter, despite his anti-transcendental gestures and rhetoric, and despite Husserl's view that Heidegger had betrayed transcendental philosophy for the sake of philosophical anthropology, Heidegger could not consistently abandon or overcome the problematic of transcendental philosophy through his displacement of the constitution of sense and meaning from the subject (*Dasein*) and its horizon of meaning to the event and openness of being (*Sein*), as advocates of his later thinking have claimed. Heidegger remained too early for being, insofar as he could not arrive at a purely ontological understanding of being and its meaning that transcended the philosophy of the subject and modernity.⁵³

Eric S. Nelson is Associate Professor of Philosophy at Hong Kong University of Science and Technology. His areas of teaching and research include modern European, East Asian, and intercultural/comparative philosophy and religion.

⁵²The most comprehensive and compelling attempt to address the issues at stake in "naturalizing Heidegger" can be found in David E. Storey, *Naturalizing Heidegger: His Confrontation with Nietzsche, His Contributions to Environmental Philosophy* (Albany: SUNY Press, 2015).

⁵³I would like to express my appreciation to the editors of this volume, Halla Kim and Steven Hoeltzel, for their encouragement, patience, and suggestions.

He is particularly interested in questions of communication, interpretation, and social interaction (hermeneutics and ethics). He has published over 70 articles and book chapters on Chinese, German, and Jewish philosophy. He is the coeditor with François Raffoul of the *Bloomsbury Companion to Heidegger* (Bloomsbury, 2013) and *Rethinking Facticity* (SUNY Press, 2008). He has also coedited, with John Drabinski, *Between Levinas and Heidegger* (SUNY Press, 2014); with G. D'Anna and H. Johach, *Anthropologie und Geschichte: Studien zu Wilhelm Dilthey aus Anlass seines 100. Todestages* (Königshausen & Neumann, 2013); and with A. Kapust and K. Still, *Addressing Levinas* (Northwestern University Press, 2005). He has a forthcoming monograph on Chinese Philosophy in early twentieth-century German thought with Bloomsbury.

9

Others as the Ground of our Existence Levinas, Løgstrup, and Transcendental Arguments in Ethics

Robert Stern

While transcendental arguments in theoretical philosophy have perhaps gone out of favor, transcendental arguments in practical philosophy have retained a reasonable following, with their proponents being more optimistic on their behalf. It is easy to see why this greater optimism is justified, for while Stroud's well-known critique of transcendental arguments in theoretical philosophy suggested that in the end they rely on a kind of idealism,¹ and while such idealism may seem problematic when it comes to the objects of our theoretical beliefs (the external world, other minds, etc.), when it comes to practical philosophy, such idealism or at least anti-realism may seem more palatable, so that here Stroud's concerns may seem to have less force. It could then be assumed that,

¹ Cf. Barry Stroud, "Transcendental Arguments," *Journal of Philosophy*, 65, no. 9 (1968): 241–56.

R. Stern (✉)

Department of Philosophy, University of Sheffield, Sheffield, UK

in practical philosophy, transcendental arguments have useful work to do, which is why they have been deployed by a diverse range of writers, from Kantians such as Christine Korsgaard, to communication theorists such as Karl-Otto Apel and his followers, to Catholic thinkers such as John Finnis. Proponents of such arguments typically take the foe to be the moral skeptic—who asks why he or she should be bound by morality and its norms—and try to show on transcendental grounds, to do with the necessary conditions for communication, or selfhood or agency or whatever, that the skeptic is always already bound by these norms—so that in some sense the skeptic’s question cannot be properly posed—and thus is answered.

My aim in this chapter is straightforward. I want to discuss first an argument of this sort that has recently been attributed to Emmanuel Levinas by Diane Perpich, as a way of understanding the former’s rather elusive views. Taking this as a representative of the use of a certain sort of transcendental argument against the skeptic, I will argue that it is not a convincing way to deal with their doubts, as it fails to achieve what it sets out to do. However, turning from Levinas to the Danish philosopher and theologian K. E. Løgstrup, I will argue that one can find transcendental reflections in his thought too, but used in a different way: not to answer the moral skeptic via a transcendental argument, but to help us show how certain fundamental misconceptions underlie moral skepticism nonetheless. I will suggest that in this role transcendental reflections can be more successful—and that perhaps this is the way we should understand Levinas’s comments too. Thus, while I think we should give up the ambition of trying to answer the moral skeptic directly using transcendental arguments, this does not mean we should abandon using some of the insights that have underpinned such arguments, in a way that reading Løgstrup can bring out.

I will begin by outlining the reading of Levinas that I want to focus on. I will then set out what I think is wrong with it as an answer to the skeptic. I will then turn to Løgstrup, to see if we can do better by adopting his approach.

But before I start, perhaps a very brief introduction to Løgstrup might be in order—for while I assume Levinas will be a familiar figure, Løgstrup is likely to be more obscure. As I have mentioned, Løgstrup (1905–81)

was a Danish philosopher and theologian. His early reading was influenced by Kant and the phenomenological movement (particularly Edmund Husserl, Max Scheler, Hans Lipps, Martin Heidegger) and Kierkegaard, as well as Lutheran theology. He spent most of his academic life at the University of Aarhus, and of course lived through the German occupation of Denmark during World War II, which had an impact on his ethical thinking in a number of ways. He published his first major work *The Ethical Demand* in 1956 (the English translation published by Notre Dame University Press appeared in 1997). He published several later books and articles in ethics, theology, philosophy of art and metaphysics (some of the later ethical writings are translated in *Beyond the Ethical Demand*, published in 2007, and a two-volume selection from the four-volume work on metaphysics was published in translation in 1995; several more works are available in German, mainly translated by his wife, whom he met while studying in Germany before the war).²

One other preliminary point to mention, but which will be more significant as we go on, is the intriguing parallels between Levinas's ideas, life, and work, and that of Løgstrup. For while it seems that no significant encounter between the two ever occurred, and certainly neither responded to the other in writing,³ they were almost contemporary (Levinas was born in 1906 and died in 1995), and shared many similar influences, where the connection with Heidegger will

² Additional bibliographical details are provided in the List of Abbreviations and in further references to Løgstrup's works below.

³ Levinas was in Strasbourg in 1923–29, then Freiburg in 1928–29, then back in Strasbourg before going to Paris in 1930 until 1940, while Løgstrup was in Strasbourg in 1930–31, Göttingen in 1931–32, Freiburg in 1933–34, and Tübingen in 1934–35. It is thus at least possible that they both attended Jean Héring's classes in Strasbourg at some point in 1930, though we have no evidence of this. Hans Hauge has also suggested via personal communication that he finds it plausible that Løgstrup might have got the idea of criticizing Husserl (in his first attempt at writing a doctoral thesis, submitted in 1933) from Levinas's own thesis defense in 1930, which was on "The Theory of Intuition in Husserl's Phenomenology." As Bjørn Rabjerg has pointed out, Husserl was almost completely unknown in Denmark at the time (there are only two references to him before Løgstrup discussed him: in 1915 and 1922 and both by the psychologist Edgar Rubin), so it is very likely that Løgstrup got this idea in Strasbourg in 1930. So perhaps not only Sartre, but also Løgstrup, "was introduced to phenomenology by Levinas," as Levinas famously observed of the former.

be important for this chapter. Both studied with Heidegger before the war, both were shocked by his political allegiances, and both wrote about him afterwards. Likewise for both, the central idea of their ethics is superficially similar—that is, while Levinas talks about the face-to-face encounter with the destitute other, Løgstrup emphasizes the ethical demand made by the other in their vulnerability and need for assistance. And both see the relation between ethics and religion in complex terms, on the one hand claiming to offer a nontheological ethics, while on the other still making use of religious idioms and thinking—in Levinas’s case the religious background is Judaism, of course, while for Løgstrup it was Lutheranism. There are thus many issues on which comparison and contrast between these two thinkers can be highly illuminating, I think—but this chapter focuses on just one, namely how claims about our prior embeddedness as subjects within a world of other individuals and their needs might play a crucial role in our reflections on the ethical and its justification.

Levinas

In his first major work, *Totality and Infinity*, Levinas begins with what many take to be a skeptical question concerning ethics: “Everyone will readily agree that it is of the highest importance to know whether we are not duped by morality.”⁴ This has been compared to what Korsgaard calls “the normative question” in her book *The Sources of Normativity*:

It is the force of . . . normative claims—the right of these concepts to give laws to us—that we want to understand.

And in ethics, the question can become urgent, for the day will come, for most of us, when what morality commands, obliges, or recommends is *hard*: that we share decisions with people whose intelligence or integrity don’t inspire our confidence; that we assume grave responsibilities

⁴ Emmanuel Levinas, *Totality and Infinity: An Essay on Exteriority*, trans. Alphonso Lingis (Pittsburgh: Duquesne University Press, 1969), 21.

to which we feel inadequate; that we sacrifice our lives, or voluntarily relinquish what makes them sweet. And then the question—*why?*—will press, and rightly so.⁵

Korsgaard's and Levinas's questions seem similar: morality requires things of us, but perhaps we are being fooled, perhaps it has no right to demand anything of us at all?

However, I think it is important to be clear about the spirit in which this question is being asked—or rather, the type of person who is asking it. One such person might be the so-called egoist of legend, who thinks the only thing she has reason to do is what is in her interests—and so when she asks this question, she wants some demonstration that morality is in her interests after all. But, as H. A. Prichard and others have argued, to try to answer this kind of skeptic—at least directly—is a mistake, because such a skeptic will only be satisfied if morality is shown to be in her interests, which might make her then *conform* to morality, but not make her *moral*, as she will then only do what is right because it is good for her to do so, and thus will never act out of moral reasons.⁶

But still, even if this direct response is hopeless and we should just not engage with the egoistic skeptic at all in this way, we could perhaps try another strategy, namely that of trying to get them to see that there are reasons to do things which are *not* in their interests, of which moral reasons would then be one. Thus, we might not try to answer the egoist directly, but *indirectly*, by providing some argument to show they have more types of reasons than they thought.

But we also might not take the normative question in this egoistic spirit at all, and thus might not take our skeptic to be an egoist. Rather, our skeptic might already perfectly well accept that we can have reason to do things that are not grounded in our interests, but *still* be skeptical about morality as we have it, with its rules and regulations that it tries to impose on us, because in fact the “morality system” is illegitimate and

⁵ Christine M. Korsgaard, *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996), 9.

⁶ Cf. H. A. Prichard, “Does Moral Philosophy Rest on a Mistake?” in *Moral Writings*, ed. Jim MacAdam (Oxford: Oxford University Press, 2002), 7–20.

can be given a “debunking” explanation, in a manner that undermines its claim to our allegiance. Thus, a figure like Bernard Mandeville (who Korsgaard mentions), or equally figures like Nietzsche or Marx or other “masters of suspicion,” try to make us see the norms of morality in a new light, in a way that shakes our confidence in them—not because we are egoists, but precisely because we are not. For example, Nietzsche puts this kind of worry with characteristic vehemence: “in so far as morality *condemns* as morality and *not* with regard to the aims and objects of life, it is a specific error with regard to which one should show no sympathy, an *idiosyncrasy of the degenerate* which has caused an unspeakable amount of harm!”⁷ In speaking of harm here, I take it, Nietzsche is not making an egoistic appeal to the harm caused to just you as an individual and expecting you to reject it as a result—rather, he is appealing to the harm it does to *us* as a whole, where it is in the failure to connect properly to our good in general that the illegitimacy of morality as it is currently practiced resides.

There is, finally, a third kind of moral skeptic, who does not doubt ethics from a purely egoistic standpoint, or claim to find in ethics some sinister facade for non-moral interests, but who just thinks that when we engage in ethics, we engage in a practice that lacks the requisite grounding to really make sense—where the most radical claim might be that ethics just can never make sense, or more moderately that it doesn’t make sense now. Thus, various kinds of nihilists, relativists, and subjectivists might count as skeptics of this radical kind—as might those who think there are grounds for denying we have the sort of freedom that arguably morality requires.⁸ And a more moderate sort of skeptic might be Elizabeth Anscombe, when she says that, while talk of moral obligation could have been intelligible when we believed in God as a moral legislator, now that we have given up that belief, our talk of moral obligation is as meaningless or empty as talk of crimes would be in a

⁷ Friedrich Nietzsche, *Twilight of the Idols*, “Morality as Anti-Nature” §6, trans. R. J. Hollingdale (Harmondsworth: Penguin, 1968), 46.

⁸ Cf. Kant’s worry that morality might be a “phantasm” in *Groundwork for the Metaphysics of Morals*, AK 4:445.

world without law courts or the police⁹—so in this sense, we are being “duped” in talking in moral language, but not because morality is being deployed for non-moral interests, but because we are being foolish in thinking we are operating within an intelligible practice when really we are not.

I do not claim that these forms of moral skepticism are exhaustive, but I do think they can be distinguished from one another. We might call the first egoistical skepticism, the second debunking skepticism, and the third metaphysical skepticism, as it questions the metaphysical framework which we seem to need to make sense of morality.

Now, I am not sure Levinas is precisely clear on which skepticism he has in mind—and this is not a question that can be gone into properly here. Instead, I now want to turn to a recent reading of Levinas by Diane Perpich, who tries to use a transcendental approach to address these issues.

Perpich begins by characterizing “the skeptic’s question” in terms of Levinas’s response to an interviewee who asked him: “What would you respond to someone who said that he did not . . . feel this call of the other, or more simply that the other left him indifferent?” Levinas, of course, thinks that this “call of the other” is central to ethics and is what it fundamentally consists in. Perpich then cites Levinas’s reply: “I do not believe that is truly possible. It is a matter here of our first experience, *the very one that constitutes us*, and which is as if the ground of our existence.”¹⁰ Perpich takes the skeptic’s challenge here to be as follows:

The skeptic asks, “What is the other to me?” or “Why ought I value the other’s demands?” Such questions imply that only a fool is duped into thinking that the other’s claims to moral consideration have any binding force apart from that already provided by utility or other prudential considerations. In effect, the skeptic doubts that there are uniquely moral reasons and doubts that such reasons have a normative force that

⁹ G. E. M. Anscombe, “Modern Moral Philosophy,” *Philosophy* 33, no. 124 (1958): 1–19.

¹⁰ Emmanuel Levinas, *Is it Righteous to Be? Interviews with Emmanuel Levinas*, ed. Jill Robbins (Stanford: Stanford University Press, 2001), 184; cited in Diane Perpich, *The Ethics of Emmanuel Levinas* (Stanford: Stanford University Press, 2008), 130, her emphasis.

cannot be reduced to self-interest. The skeptic thus asks what reasons there are to value something or someone she has been told she has an obligation to value. She asks why she should think that the other has *moral* value or deserves moral consideration.¹¹

Thus, of the three forms of moral skepticism outlined above, it is the egoistic skeptic that Perpich sees as Levinas's concern here.¹² And taking Levinas's interview response a certain way, Perpich glosses his answer as follows:¹³

Faced with this ordinary sort of moral skepticism, a Levinasian response might begin like this: the idea of valuing something presupposes a world in which this something is meaningful or intelligible. If I value something—regardless of the value I assign to it—this implies I have taken its measure, weighed it against other objects or possibilities, and in some manner understood its connection to my own life and to others . . . Valuing requires and expresses the fact that I am *already reflectively in a world* . . . Being in a world presupposes an other who has *opened that world to me and with me*. I do not meet the other *in* the world; rather, to have a world (which means being capable of reflection) is already to be in a relationship to the other. Without

¹¹ Perpich, *Ethics of Emmanuel Levinas*, 130–1.

¹² And also at the beginning of *Totality and Infinity* in the passage cited, where Perpich notes “that Levinas imagines the skeptic in this way is evident in the opening lines of *Totality and Infinity*.” Perpich, *Ethics of Emmanuel Levinas*, 209 n. 5.

¹³ Perpich notes that her way of presenting things “illuminates [Levinas’s] position from a different perspective and in a different idiom,” but still claims that it “illuminates the structure of his thought” (*Ethics of Emmanuel Levinas*, 209 n. 7; cf. 134–5 and 14). Michael Morgan also offers what he calls a “transcendental reading” of Levinas, but without seeming to identify any transcendental argument as such: see Michael L. Morgan, *Discovering Levinas* (Cambridge: Cambridge University Press, 2007), 50–60. Transcendental aspects of Levinas’s thinking are also discussed and debated in: Robert Bernasconi, “Rereading *Totality and Infinity*,” in *The Question of the Other: Essays in Contemporary Continental Philosophy*, ed. Arleen B. Dallery and Charles E. Scott (Albany: SUNY Press, 1989), 23–34; Theodore de Boer, “An Ethical Transcendental Philosophy,” in *Face to Face with Levinas*, ed. Richard A. Cohen (Albany: SUNY Press, 1986), 83–115; Jeffrey Dudiak, *The Intrigue of Ethics: A Reading of the Idea of Discourse in the Thought of Emmanuel Levinas* (New York: Fordham University Press, 2001); and Brian Treanor, *Aspects of Alterity: Levinas, Marcel and Contemporary Debate* (New York: Fordham University Press, 2007).

the other, there is no world; without the world, there is no ego who could be the subject or bearer of experiences within the world. The relation to the other is thus constitutive of my having a world at all.¹⁴

Now, Perpich does not use the term “transcendental argument” here or elsewhere in the book, but I take it that her reading of Levinas’s position is meant to have a transcendental flavor, in relation to two claims she makes here. First, that to value anything requires a reflective process of some kind; and second that there is a constitutive relation between the self and the other, where therefore both claim that one thing is a necessary condition for the possibility of another—reflection is a necessary condition for valuing, and intersubjectivity is a necessary condition for selfhood. Suggestions of this sort are characteristic of transcendental arguments. Moreover, so is the way she deploys these claims, and the way she tries to use them against the skeptic:

[The argument] now reads in full as follows: (1) The idea of holding anything as a value presupposes a world in which the thing valued is already meaningful or intelligible. To value thus means I am already reflectively situated in a world. (2) Being reflectively in the world is the product of a social relationship, which is to say that the relation to the other is constitutive of my being able to value anything whatsoever. (3) Thus, the other person is not merely something or someone that I can value or fail to value. Without the other, there is no world and no meaningful valuing. Hence the skeptic’s question of whether I am obliged to value the other and on what grounds always comes *too late*. It mistakes the other for an object within the world, rather than seeing the relationship to the other as the condition of my having a world at all and being able to find value in it. If I value *anything* at all, then, I am already in a relationship to the other. He or she *already* concerns me. What could my continuing to ask for proof of this mean except that I have failed to understand what sort of relationship we have? The other can never be only an object of value within the world concerning which I might rightfully ask why she or her needs should matter to me. By the time I ask these questions, I have already shown myself to be

¹⁴ Perpich, *Ethics of Emmanuel Levinas*, 131–3.

immersed in a complex evaluative practice. In effect, to ask for reasons, to ask why I should concern myself with the other, is itself already indicative of such concern. The skeptic's question thus indicates that the other has already passed that way, already introduced her into a world in which critical reflection is possible.¹⁵

Here, to adapt slightly Kant's well-known remark concerning the Refutation of Idealism, the "game played by the sceptic has been turned against itself"¹⁶ in a way that many take to be distinctive of transcendental arguments, where the skeptic is shown to be guilty of what Apel famously calls a "performative contradiction." For, Perpich thinks, what this Levinasian argument shows is that the skeptic, in order to be a subject at all, must already have been embedded in an ethical or social relation to an other, so that in questioning those relations she must already have shown concern towards them, so the question comes "too late."

Perpich puts this point as follows:

From this vantage point, the skeptic's question is put in a new light. The would-be amoralist asks for proof or evidence that the other is his concern: "What is my brother to me or I to him that I should concern myself with his welfare?"¹⁷ The skeptic effectively demands a *reason* that would justify the other's demand for care or concern. In so doing, the skeptic implicates herself in the very practices of reflection that indicate just the sort of relation she would like to deny. That is, the skeptic uses a faculty or practice granted to her by the social or ethical relationship in order to

¹⁵ Perpich, *Ethics of Emmanuel Levinas*, 134.

¹⁶ B276. Kant of course talks about the idealist, rather than the skeptic.

¹⁷ Perpich is here referring to a passage in *Otherwise than Being*, where Levinas writes:

Why does the other concern me? What is Hecuba to me? Am I my brother's keeper? These questions have meaning only if one has already presupposed that the ego is concerned only with itself, is only a concern for itself. In this hypothesis it indeed remains incomprehensible that the absolute ego posited for itself speaks a responsibility. The self is through and through a hostage, older than the ego, prior to principles. What is at stake for the self, in its being, is not to be. Beyond egoism and altruism it is the religiosity of the self.

Emmanuel Levinas, *Otherwise than Being, or Beyond Essence*, trans. Alfonso Lingis (Pittsburgh: Duquesne University Press, 1998), 117. Cf. Perpich, *Ethics of Emmanuel Levinas*, 135.

question whether such a relation could really be attributed to her. Her question thus involves her in a performative contradiction and is in this sense self-defeating or self-refuting.¹⁸

Perpich goes on:

In demanding a justification, the moral skeptic is trapped in a performative contradiction between the content of her question and its practical conditions. The skeptic's question presumes a neutral, pre-social subject who has no constitutive relation to the other and thus must be provided with reason to take the other into account. But the practices of reasoning in which the skeptic's own question participates already belie her introduction into a socially or intersubjectively constituted world. When the skeptic asks "Why be moral?" or "What is the other to me?" she demands a reason for acting in one way rather than another. Far from casting doubt on the possibility of ethical life through such questions, skepticism is in fact its prolongation. It is the *enactment* of ethical life. If it were not for the other who opens the world to me, I would not be able meaningfully to ask the skeptic's question. Thus, being chosen before I can choose is the condition for all of my later choosing, for all my affirming or denying. I cannot without contradiction deny my ability to engage in the process of critical reflection, and, by extension, I cannot without contradiction deny my exposure to the other. This inability to turn a deaf ear, this non-indifference to the other, *is* the moment of normativity in Levinas's thought.¹⁹

Set out as an argument, therefore, I think the structure of the performative contradiction Perpich has in mind is as follows:

1. To ask "why concern myself with others?" is to engage in critical reflection.
2. To engage in critical reflection is to be a subject.
3. To be a subject it is necessary to be concerned with others.

¹⁸ Perpich, *Ethics of Emmanuel Levinas*, 143–4.

¹⁹ Perpich, *Ethics of Emmanuel Levinas*, 145–6.

4. So to ask “why concern myself with others?” I must already have concerned myself with them.
5. So to ask “why concern myself with others?” is performatively contradictory.

I now want to raise some doubts about this argument, where I will grant premises 1 and 2, but focus on the transcendental claim in premise 3, and worry about whether the final conclusion in 5 really resolves the skeptical issue from which Perpich starts.

Regarding premise 3, the question is this: suppose the egoist accepts that she could not be a subject at all unless she showed concern to others—how will that satisfy her? One way is that it might give her an egoistic reason to be concerned with others: that is, she might accept that if she could not exist at all unless she showed concern for others, she has every reason to show such concern. But this is presumably hardly a satisfying result to the moralist, as it now just gives the egoist a non-moral reason to be moral, as Prichard and others have feared.

However, it might be said, this is mistakenly to take premise 3 *on its own* to answer the skeptic, but the point is to *use* premise 3 to argue for the performative contradiction in the conclusion, where it is *this* that is meant to answer the skeptic, by showing that she cannot properly doubt whether she should concern herself with others, as in doing so she must presuppose such concern has gone on, as otherwise she would not be a subject at all who can even ask this question. As Perpich puts it in a passage we have already cited: “if I value *anything* at all, then, I am already in a relationship to the other. He or she *already* concerns me.”²⁰

I am not convinced, however, that this kind of strategy (and indeed performative or retorsive transcendental arguments in general, to use the other term that is sometimes applied to them) can really resolve the issue.²¹ For it seems to me that the skeptic can reply by saying that, even

²⁰ Perpich, *Ethics of Emmanuel Levinas*, 134.

²¹ For further discussion, see my “Silencing the Sceptic?” in *Transcendental Arguments in Moral Theory*, ed. Jens Peter Brune, Robert Stern, and Micha Werner (Berlin: de Gruyter, forthcoming). I here distinguish performative or retorsive transcendental arguments which try to convict skeptics of undercutting their own position by denying what they must presuppose, from deductive

if Perpich's transcendental claim in 3 is right, all this leads to is the conclusion that, to be a subject at all, I must have gone in for concern for others—but what the skeptic wants to know is whether she was *right* or *justified* in doing so, or whether to do so was misguided and she was being “duped.” The fact that she *has to have been concerned* in order to be a subject in the first place would not seem to resolve this worry.

Now, one response might be to go back to the thought that justification for the concern can come from the idea that, without this concern, I wouldn't be a subject at all—so I have excellent egoistic reasons for thinking I was right to be concerned, and far from being duped, I have done the right thing. But again, this is to give a grounding for concern that is merely egoistical in nature, which I take is inadequate to the moralist, who is looking for a grounding which will show the egoist why she *should not* be an egoist, rather than giving her egoistic reasons to be moral.

Another response might be that, in a situation where some form of activity—in this case, concern for others—is necessary, to ask for reasons for engaging in it is idle: if, as a subject, I *must* be concerned for others because otherwise I couldn't be a subject at all, then that is all the justification that is needed, as it shows I have no other option but to be concerned in this way. On its own, however, this seems unpersuasive, because it seems we can ask normative questions regarding features of ourselves and our lives that we see to be necessary, even metaphysically so—for example, death is necessary to creatures such as us, but we can still ask whether it is good or bad to die.

However, this may seem to miss the point; for it could be said that what makes this a performative contradiction is that it is not just necessary to us in a general way, but necessary to being a subject and thus even raising the question at all, which makes it something we must take for granted in all subsequent reasoning—thus, to use a familiar metaphor, on this view the question “why concern myself with others?”

transcendental arguments which try to establish the falsity of the skeptics' position on the basis of premises that they accept via the claim that the former is a necessary condition of the latter holding true. I also argue that the promise of performative or retorsive transcendental arguments is illusory.

is an “external” question rather than an “internal” one, and as such does not even require an answer, any more than the question “is the principle of non-contradiction valid?” In even raising the question, we are always already committed to this concern, because we could not ask the question at all if we were not. This then explains why I am *obliged* or bound to be concerned by the other: as a subject, I have no alternative option, so it is not possible to cast doubt on this obligation in the way the skeptic tries to do, thereby silencing her. As Perpich puts it: “without the other, there is no world and no meaningful valuing. Hence the skeptic’s question of whether I am obliged to value the other and on what grounds always comes *too late*.”²²

But here we meet a familiar worry about such arguments: namely that even if they show that believing or acting in a certain way is some sort of necessary presupposition, this doesn’t in itself give a reason to think such presuppositions are true or justified²³—and because it doesn’t, the fact that the argument seems to show that such presuppositions cannot be given up or doubted seems to make the skeptical challenge *worse*, not better, as now we seem to be compelled to believe or act without being supplied with any grounds to justify these beliefs or actions. So, the skeptic might worry, all I have been shown is that there is no way qua subject not to be concerned by others—but I still haven’t been shown I have grounds for such concern that would merit it, and thus that I am not being duped; all I have really been shown is that I am in a situation

²² Perpich, *Ethics of Emmanuel Levinas*, 134.

²³ Cf. Charles Sanders Peirce, *Collected Papers*, 8 vols, ed. Charles Hartshorne, Paul Weiss and Arthur W. Burks (Cambridge, MA: Harvard University Press, 1931–60), vol. 2, §113:

I do not admit that indispensability is any ground of belief. It may be indispensable that I should have \$500 in the bank—because I have given checks to that amount. But I have never found that the indispensability directly affected my balance, in the least . . . When we discuss a vexed question, we hope that there is some ascertainable truth about it, and that the discussion is not to go on forever and to no purpose. A transcendentalist would claim that it is an indispensable “presupposition” that there is an ascertainable true answer to every intelligible question. I used to talk like that, myself; for when I was a babe in philosophy my bottle was filled from the udders of Kant. But by this time I have come to want something more substantial.

where I cannot even question this, which may deepen my skepticism rather than alleviate it. Moreover, as an answer to the question of obligation, the argument seems misconceived, for while it may show that I am obligated to concern myself with others in the sense of being unable to do otherwise, it does not show that others have a *legitimate authority* over me and hence obligate me qua rational agent, which is surely what the skeptic is asking to be shown.

Finally, however, it might be said that Perpich's argument can be made stronger than I have presented it so far, if instead of an argument focused on concern for the other, we rather focus it on the *value* of the other—for if we can show the egoist that the other has value, then this might seem to give her a *reason* to feel concern, which we have been missing so far. Thus, we might gloss Perpich's argument as follows:

- (1) To ask “why concern myself with others?” is to engage in critical reflection.
- (2) To engage in critical reflection is to be a subject.
- (3') To be a subject it is necessary to be in an ethical and social relation with others.
- (4') To be in a social relation with others it is necessary to value them.
- (5') So I must take others to be valuable.
- (6') So concern for others is justified.

In this way, it could be argued, the skeptic's demand for justification that drives the normative question is answered.

However, of course, the difficulty here is with premises (4') and (5'), which again involve the sense in which it is necessary to value others, and what this means. For, of course, one way one might feel that morality could “dupe” us is if it turned out we were forced into valuing things in ways that were not merited—and by demonstrating merely that one *must* value others, it is hard to see how this worry would be allayed. Of course, if one is sufficiently idealist about value, one might argue that what it is to *be* valuable is to have value conferred on it by our attitudes and preferences—so if we must have a valuing attitude to X, then X ipso facto possesses value, and to think otherwise is to ask for more than is required. But then, the value of others can no longer serve as the reason for our social and ethical

relations with others: rather, it is itself constructed through those relations. However, this would then leave the skeptic who wonders about the reason for such relations without any answer based on the value of others; the reason would only seem to be that such relations are necessary to be a subject at all, thus taking us back to the earlier form of the argument, and the difficulties we identified there.

It would seem, then, that, notwithstanding Perpich's undoubtedly very interesting reconstruction of Levinas's position, any insights it may possess still fail when taken as a transcendental argument against the moral skeptic.

Løgstrup

Having seen how I think the approach outlined above fails, I now want to turn to Løgstrup, to contrast the way he uses somewhat similar claims and insights, but in a different dialectic, which is arguably more successful.

We might begin with the similarities between Levinas and Løgstrup. I think an important basis for this similarity is their shared background in Heidegger, which has been noted. When it comes to Levinas, I think Perpich is right to highlight the indebtedness to Heidegger in his key claim, namely that as subjects we are not "pre-social" beings who reside outside the social context and choose whether to engage with it (or not), but that we are always already embedded within it from the start:

To return to our Levinasian argument then: valuing requires and expresses the fact that I am *already reflectively in a world*. To be in-the-world can be understood here with the full richness Heidegger gives to the term. I am immersed in an open-ended system of relationships, many of which I understand and control, some of which I do not, all of which refer to possibilities of the kind of being that I am myself. What are we to say about how I came to be there? While Heidegger takes the question of the world-hood of the world to be one of the fundamental questions of ontology, he would no doubt read the question of *how* we come to find ourselves in a world as an ontic question of little direct interest. On the Levinasian account, however, how we answer the question is crucial. Being-in-the-world, for

Levinas, is neither the achievement of a self-sufficient subject nor the ontological birthright of *Dasein*. Being in a world presupposes an other who has *opened that world to me and with me*.²⁴

As others have argued,²⁵ it is possible to give these aspects of Heidegger's thought an ethical construal—even if he resisted any such construal himself—and arguably this forms an important background to Levinas's thinking here.²⁶

Likewise, I think the same is true of Løgstrup, for he also emphasizes the way in which we are not sovereign when it comes to our own lives but depend fundamentally on others through relations of trust, communication, and care, and equally that, when we act in accordance with norms like trust and openness to others, we are not following norms that we have somehow created for ourselves as sovereign individuals—rather, seen as the norms which structure our lives from the outset, they might be said to have a certain degree of sovereignty over us, which is why Løgstrup calls them in his later works “sovereign expressions of life,” and in his earlier works speaks of them as a “gift” in the sense that they are “givens.”²⁷ Thus, for example, in *The Ethical Demand*, Løgstrup writes as follows concerning trust:

It was said earlier that a person does not arbitrarily deliver himself over to someone else as a matter of trust. Rather, this self-surrender is a part of his life, irrespective of any decision on his part. Also it was said that this

²⁴ Perpich, *Ethics of Emmanuel Levinas*, 132.

²⁵ See for example Frederick A. Olafson, *Heidegger and the Ground of Ethics: A Study of Mitsein* (Cambridge: Cambridge University Press, 1998).

²⁶ This is not to say that he does so without any reservations: cf. Levinas, *Is it Righteous to Be?*, 177: “in Heidegger, the ethical relation, *Miteinandersein*, being-with-another, is only one moment of our presence in the world. It does not have the central place. *Mit* is always being next to . . . It is not in the first instance the face, it is *zusammensein* (being-together), perhaps *zusammenmarschieren* (marching-together).” Cf. also 137.

²⁷ What precisely Løgstrup means when he speaks of life being a gift in *The Ethical Demand* is somewhat controversial, but I am here following the suggestion of Hans Fink and Alasdair MacIntyre that his argument relies on the idea of “life being something *given* in the ordinary philosophical sense of being prior to and a precondition of all we may think and do.” Hans Fink and Alasdair MacIntyre, “Introduction,” in ED xxxv.

implies the demand that we take care of the life which has been placed in our hands . . . That a person is more or less in the power of another is a fact we cannot alter; it is a fact of life. We do not deliberately choose to trust, and thereby deliver ourselves over to another. We constantly live in a state of being already delivered—either through a passing mood or in terms of something which in a fundamental way affects our entire destiny.²⁸

For Løgstrup, relations of trust are not practices that people institute for themselves as individuals. Rather, as what Alasdair MacIntyre has called “dependent rational animals,”²⁹ we are constantly caught up in relations of trust, not only for information but in the very process of communication itself, which requires what Løgstrup calls “openness of speech”—that is, honesty in our engagements with one another. Of course, this does not mean that *all* people are to be trusted, that *all* speech is open, all the time. But it does mean (Løgstrup thinks) that unless this were the norm, then human life as we know it would not be possible, making it in this sense a transcendental condition for such life.

As we as individuals are subjects embedded in this life, rather than sovereign individuals who somehow create ourselves from scratch, there are clear echoes in Løgstrup of the key thought that Perpich attributes to Levinas, namely that we could not be subjects at all without this ethical and social context in which we exist, so we are constitutively dependent on others in this sense. Thus, Løgstrup writes:

It is not within our power to determine whether we wish to live in responsible relations or not; we find ourselves in them simply because we exist. We are already responsible, always, whether or not we want to be, because we have not ourselves ordered our own lives. We are born into a life that is already ordered in a very definite way, and this order lays claim upon us in such a manner that as we grow we find ourselves bound to other people and forced into responsible relationships with them.³⁰

²⁸ ED 54–5.

²⁹ Cf. Alasdair MacIntyre, *Dependent Rational Animals: Why Human Beings Need the Virtues* (London: Duckworth, 2009).

³⁰ ED 107.

In the following passage, therefore, Løgstrup explicitly treats the normative structure under which we live, including trust and care for others, as constitutive of our identity:

[The sovereign expressions of life] reach back into and are given with something for which we are not ourselves the basis. We cannot reserve them for ourselves as our contribution or society's, pleading that they are personal. Yet, we are on intimate terms with them. Our identity is literally due to them. They constitute it. Our identity is due to life-expressions over which we have no power except that we can annihilate them. They reach beyond the regional compartmentalization upon which we otherwise depend.³¹

Like Levinas, though without mentioning Heidegger explicitly, there is a Heideggerian element to Løgstrup's thinking here that perhaps explains their common ground. Thus, both can be read as offering a kind of transcendental claim about our necessary prior embeddedness as socially and normatively structured beings, on which our individuality depends.

However, while this much is shared between Levinas (as presented by Perpich) and Løgstrup, there is also an important difference, which is that, while Levinas is put forward as using these claims against the skeptic in a transcendental argument, to try to convict the skeptic of performative contradiction, Løgstrup seems to have no such ambition. He makes no attempt to defeat or silence the skeptic using an argument of this sort as such, and nor does he write as if such a strategy is required. Thus, despite sharing many premises and assumptions, there is nothing in Løgstrup to mirror the sorts of transcendental arguments we have sketched in relation to Perpich's reading of Levinas.

However, it might now be puzzling why Løgstrup made a point of underlining our condition as socially and ethically embedded individuals, if his aim was not to construct a transcendental argument of a Levinasian sort. For, it could be asked, *unless* we are using these claims

³¹M1 88, translation modified. Cf. also ED 66: "to be an individuality, a self, implies that something is claimed of me."

against the moral skeptic in a transcendental argument, what is the point of dwelling on them? What useful work could they be made to do otherwise? Isn't their only point to silence the skeptic, otherwise these insights are wasted? It is these assumptions that I think Løgstrup's writings usefully challenge, by showing that there is another (perhaps ultimately more successful) way in which such transcendental insights can be used.³²

To see this, it is significant first to note the dialectical situation in which Løgstrup thinks he is operating, given that he thinks it does not need to be met by a transcendental argument: what kind of challenge is it, instead? If the challenge is coming from the moral skeptic, won't it need a transcendental argument to address it properly? And if it is not coming from the moral skeptic, how can we take Løgstrup and Levinas to be engaged in the same endeavor at all? To answer these questions requires some care.

As a first step, I think it is useful to go back to Perpich's starting point in her construction of Levinas's transcendental argument, when she cites a question he was asked by an interviewer: "What would you respond to someone who said that he did not . . . feel this call of the other, or more simply that the other left him indifferent?" before providing Levinas's reply: "I do not believe that is truly possible. It is a matter here of our first experience, *the very one that constitutes us*, and which is as if the ground of our experience." As we have seen, Perpich thinks that what concerns Levinas's interviewer here, and Levinas himself, is whether there are genuine reasons to care for others, to which the performative contradiction involved in doubting this is supposed to offer some sort of response.

But I think a different reading of Levinas's exchange with his interviewer is possible. For what is notable is that the latter raises a question about what someone *feels*, where what he asks Levinas is whether he thinks anyone could be oblivious to the other and be left indifferent to

³² For more on the distinction between transcendental arguments and transcendental claims that is developed in what follows, see Robert Stern, "Taylor, Transcendental Arguments, and Taylor on Consciousness," *Hegel Bulletin* 34, no. 1 (2013): 79–97.

him or her.³³ Now, this looks more like a *phenomenological* question than a question about *reasons*: it seems as if Levinas is being asked to comment on whether he thinks a person could fail to register the call of the other in his or her experience of the world, not whether a person could see that call as failing to provide a reason to act, which would be a further step. Likewise, Levinas's response seems not to be to claim that no one could see the call as a reason to act, but rather that no one could fail to hear the call in the first place, whether or not he or she then takes it to be reason-giving. What seems "not to be truly possible" because it is "a matter here of our first experience" is to fail to feel the call of the other, where this feeling is said to be "a matter of our first experience" because it is this feeling that "first constitutes us." Read phenomenologically (contra Perpich), therefore, the transcendental claim about what constitutes us is not meant to provide us with a reason to heed the call of the other; it is just meant to give grounds for thinking that everyone, including the skeptic who claims to feel indifferent, must at some level feel this call, in order to be a subject at all.³⁴

But even if this is perhaps a more accurate reading of this exchange between Levinas and his interviewer, it might still be wondered whether it makes much ultimate difference to what is going on. For even if we can convince the skeptic that she must at least feel the call of the other, can't she still ask why she should take it seriously—why not ignore it? Surely in order to address this question, we need reasons why the call is valid or legitimate, so won't something more like Perpich's argument be required after all?

However, the phenomenological reader might reply: this is to misunderstand what is meant here by "call." For, it could be argued, to be called

³³ Cf. Gary Gutting, who observes that "Levinas's response to ethical skepticism appeals to experience rather than to argument or linguistic analysis." *Thinking the Impossible: French Philosophy since 1960* (Oxford: Oxford University Press, 2011), 125.

³⁴ Cf. *Is It Righteous to Be?*, 50, where Levinas responds to the question "How to encounter the other?" by saying: "To encounter, what does that mean? From the very start you are not indifferent to the other. From the very start you are not alone! Even if you adopt an attitude of indifference you are obliged to adopt it! The other counts for you; you answer him as much as he addresses himself to you; he concerns you!" And for a more general comment that may be relevant here, cf. p. 221: "In reflecting on the transcendental conditions of the poem, you have already lost the poem."

by the other is not just (as it were) to hear some noises, which one may or may not decide to take as reason-giving—as when I hear a creak on the stairs at night, and wonder if it gives me reason to call the police, or reason to go back to sleep but call the carpenter in the morning to get my floorboards fixed. Rather, Levinas would surely hold, to feel or experience the other as requiring something of us precisely is to see what is going on *as* reason-giving, as otherwise one hasn't really heard it as a call of the other at all—just as something unconnected to him or her in any way.³⁵

Nonetheless, Perpich might say, questions concerning reasons can still come back in. For, it could be argued, even if Levinas on this phenomenological reading were right to say that everyone in some sense experiences the call or demand of the other on them, and even if this must involve some *prima facie* sense in which they are being given a reason to act, we can still surely ask whether that *prima facie* sense is accurate, and whether we are *really* being given a reason—just as, knowing that I am addicted, I might wonder whether the call of the heroin to me to take it is *really* giving me a reason to do so. And faced with this question, Perpich might again suggest, convicting the skeptic of performative contradiction in asking it could ensure that it gets a positive answer, by showing that we must value others in such a way as to show we have a genuine reason to heed the call.

However, if my previous discussion was right, unfortunately this is not how things turned out: the transcendental argument did not really

³⁵ Cf. Korsgaard, *Sources of Normativity*, 140:

If I call out your name, I make you stop in your tracks. (If you love me, I make you come running.) Now you cannot proceed as you did before. Oh, you can proceed, all right, but not just as you did before. For now if you walk on, you will be ignoring me and slighting me. It will probably be difficult for you, and you will have to muster a certain active resistance, a sense of rebellion. But why should you have to rebel against me? It is because I am a law to you. By calling out your name, I have obligated you. I have given you a reason to stop.

Cf. the interview with Levinas cited previously, where he continues: “however indifferent one might claim to be, it is not possible to pass a face by without greeting it, or without saying to oneself, ‘What will he ask of me?’ Not only our personal life, but also all of civilization is founded upon this” (*Is it Righteous to Be?*, 184). Cf. *ibid.*, 216: “when you have encountered a human being, you cannot simply leave him alone.”

seem to supply the right kind of reasons, so this is where we may seem to have reached an impasse. On the one hand, Levinas may have established that no one can fail to experience the call of the other; on the other hand, Perpich's transcendental argument seems to have failed to allay doubts the skeptic may have concerning whether or not she should heed the call, or if in doing so she would be being "duped."

However, a way out of this impasse can be found, I think, if we consider an issue which we have left aside up to now, but which Løgstrup can help us see more clearly: namely, we have not yet considered the *grounds* on which the skeptic might question the call of the other. Assuming with the phenomenological argument that no one fails to experience this call because it is a transcendental condition on being a subject at all, but also assuming that there is no satisfactory transcendental argument to show that this alone puts it beyond questioning, nonetheless if we are to take that skepticism seriously, then we still need to be convinced that there are *legitimate grounds* the skeptic has for such questioning, and shown what they might be.

Now, Løgstrup suggests, when we come to think we can rightfully reject the demand of the other, this is usually because we precisely forget the Heideggerian claim about embeddedness, and so easily fall into thinking of *ourselves* as sovereign in ways that we really are not, where we take this to warrant us in rejecting the call that others seem to make on us. This is something that Løgstrup warns us about repeatedly:

We live in yet another comprehensive illusion. This consists of thinking, feeling and acting as if we ourselves were the power to exist in our existence. In every way, we conduct ourselves as if we owed our existence to ourselves.³⁶

As already indicated, we live in a fundamental delusion in everything we think, feel and do which is just as grotesque as it is self-evident—namely that we owe our existence to ourselves. It is not a delusion which overpowers only occasionally but one in which we live and breathe, spiritually speaking.³⁷

³⁶ M1 72.

³⁷ M1 91.

If a person refuses to acknowledge any demand for unselfishness, he or she thereby also refuses to acknowledge that his or her life has been received as a gift.³⁸

Løgstrup's thought might be put this way: while we all hear the ethical demand, we block its reason-giving force by taking it that "we owe our existence to ourselves" as somehow self-created, sovereign individuals, to whom the demand does not apply—which is how we then come to ask the questions that concern Perpich, namely "What is the other to me?" or "Why ought I value the other's demands?"

However, rather than treating these questions (as Perpich does) as entirely abstract questions, as based simply on denying the *prima facie* reason-giving force of the other just as such, which means they can only be met by a transcendental argument or not at all, Løgstrup's approach implies that these questions are only ones to be taken seriously insofar as they have some real ground based on assumptions about our sovereignty; this then means that transcendental considerations that show how false this appeal to sovereignty is can be applied at this level instead—not as premises in a transcendental argument directed against a groundless skepticism, but as a transcendental claim concerning how we relate to others which can then undercut the basis on which the skeptic tries to block the moral demand.

Thus, I have suggested, it is a mistake to use the transcendental claim that we are necessarily constituted by others to block directly moral skepticism by trying to convict the skeptic of performative contradiction, as Perpich (and others) try to do. Rather, I have suggested, it is better to use this claim in a different way, to undermine the skeptic's reasons for thinking the claims of others do not apply as reasons to her, insofar as she is a sovereign individual who thinks "[she] is lord and creator of [her] own life, in other words, that [she] has received nothing."³⁹ What

³⁸ ED 116.

³⁹ ED 156.

the Heideggerian considerations that both Levinas and Løgstrup appeal to show, therefore, is not only that we must hear the call of the other in so far as we are always already within a social world, but also that this sense of our own sovereignty that we use to block the call is a delusion, for such sovereignty is impossible for us. But what our reading of Løgstrup has likewise suggested is that certainly he, and perhaps also Levinas, would not have wanted to go further and offer a direct transcendental argument against moral skepticism based on such claims.

However, a final worry might now be raised: namely, has anything I have said about Løgstrup's argument (and Levinas's on this second interpretation) really done enough to answer the moral skeptic; and if it hasn't, isn't Perpich's approach to be preferred, at least in having the right kind of ambition? For, it could be argued, the moral skeptic is an egoist, who thinks only her interests give her reasons to act; thus, even if she may hear the call of the other, she will think it gives her no reason to heed the call unless doing so will serve her ends. This is, as it were, a theory about practical reasoning. But what do Løgstrup's Heideggerian observations do to block this theory or show it to be false? Even if Løgstrup is right to emphasize that we are not "worldless individuals, ourselves the authors of our goals—as though there were not a challenge that proceeds to us from the world and its order,"⁴⁰ why can't the skeptic stick to her account of practical reasoning and so reject as illusory the reasons this world seems to present to us? To address this challenge, won't something more like Perpich's transcendental argument be required?

Here, I think, it is necessary to return to the distinction between kinds of skepticism that I outlined at the beginning. For Løgstrup, and I think also for Levinas, the skeptic that interests them is not (as it were) the pure moral skeptic with her egoistic model of practical reasoning: for very few of us are like that, making the model something of a

⁴⁰ K. E. Løgstrup, *Norm og spontaneitet* (1972), partially translated in *Beyond the Ethical Demand*, ed. Kees van Kooten Niekerk (Notre Dame: University of Notre Dame Press, 2007), 95.

philosophical construct.⁴¹ Rather, as the opening of *Totality and Infinity* suggests, we become moral skeptics not because we start by thinking only our interests give us reasons, but because we think the other kinds of reasons that are presented to us lack sufficient substance in their own right, as the “debunking” skeptic suggests. Thus, faced with the claim of the other, we do not reject it on the grounds that, while the other is perfectly entitled to make the claim, it is nonetheless only a reason for us to act if it coincides with our interests. Rather, we question the entitlement of the claim, on the grounds that we are not responsible for the suffering of the other, or that it can only be legitimate if it involves reciprocity, or that our lives are to be left at our own disposal in so far as we deserve credit for them. It is to these “moralized” objections to the call of the other that Løgstrup’s argument is addressed, by showing that they would only apply to sovereign individuals which we are not and cannot be, and thus cannot be used to block the claim others make upon us. If this skeptical target is less radical than the moral egoist in a way that may seem to devalue this response, Løgstrup would also, I think, take it to be more real, and thus it is ultimately more important to our moral lives that this is the target that is addressed.⁴²

Robert Stern has been Professor of Philosophy at the University of Sheffield since 1989. His main interests in the history of philosophy are in nineteenth-century post-Kantian German philosophy, especially Hegel. His interests in contemporary philosophy are in epistemology, metaphysics and ethics. He is currently working on the Danish philosopher and theologian K. E. Løgstrup. His first book was *Hegel, Kant and the Structure of the Object* (Routledge, 1990), and he has also written monographs on *Transcendental Arguments and*

⁴¹ That is one way of putting it. Another, suggested by some of Levinas’s comments, is that in *some* sense the egoist is just right, and there is a kind of “madness” or “irrationality” in responding to the ethical demand—but still not in a way that need trouble us, or that requires some transcendental argument in response. Cf. *Is it Righteous to Be?*, 145: “Now let us approach something truly mad: I must care for *your* being. I cannot allow myself to abandon you to your death. This madness is what is human.” See also 190: “[Charity] is wisdom, which interrupts the good sense of the interested animal”; and 250, where he calls our responsibility for the other “madness in a way.”

⁴² I am grateful to Diane Perpich for her helpful comments on a previous draft of this chapter, and also to the editors of this volume.

Scepticism (Oxford University Press, 2000) and *Understanding Moral Obligation: Kant, Hegel, Kierkegaard* (Cambridge University Press, 2012), as well as a commentary on *Hegel's Phenomenology of Spirit* (Routledge, 2002). Two collections of his papers have been published: *Hegelian Metaphysics* (Oxford University Press, 2009) and *Kantian Ethics* (Oxford University Press, 2015). He has also edited volumes on Hegel, transcendental arguments, and pragmatism, and run a three-year research project on idealism and pragmatism.

10

Raising Validity Claims for Reasons Transcendental Reflection in Apel's Argumentative Discourse

Matthias Kettner

Karl-Otto Apel, contemporary German philosopher, is best known for his wide-ranging philosophical approach of a “transcendental pragmatics of communication” (“*Transzendente Sprach-Pragmatik*,” TPC). TPC has profound implications for a gamut of issues in theoretical and practical philosophy that would amount to a “transformation” of foundationalist projects that have been elaborated in the older tradition of ontological metaphysics and had already been thoroughly transformed with the elaboration of Kant’s new paradigm of transcendental reflection.¹ Apel’s

¹ Cf. Karl-Otto Apel, “What is Philosophy? The Philosophical Point of View after the End of Dogmatic Metaphysics,” in *What is Philosophy?*, ed. C. P. Ragland and Sarah Heidt (New Haven: Yale University Press, 2001), 153–82. For Apel’s theory of paradigms of first philosophy, see: Karl-Otto Apel, “Transcendental Semiotics and the Paradigms of First Philosophy,” in *From a Transcendental-Semiotic Point of View*, ed. Marianna Papastephanou (Manchester, Manchester

M. Kettner (✉)

Department of Philosophy, University of Witten/Herdecke, Witten,
Germany

e-mail: kettner@uni-wh.de

intention was to defend and revise Kant's paradigmatic shift from ontological to transcendental reflection by introducing a communicative dimension to transcendental reflection. Other philosophical resources that have shaped TPC are semiotic pragmatism along Peircean lines, Wittgenstein's pragmatism of language games, and continental traditions of the philosophy of language and of hermeneutics from Humboldt to Heidegger.²

TPC accords "argumentative discourse" and its essential normative presuppositions a foundational as well as critical role within all other philosophical inquiries in which justifiable validity claims are raised (for example, in epistemology, normative theories of rationality, and ethics). If there are such presuppositions, then any interlocutor's communicative intention to waive them, when involved in an instance of argumentative discourse, will clash with the construal of that discourse as rationally meaningful, since it involves the interlocutor in a philosophically interesting kind of inconsistency. For Apel (and for Habermas), this kind of pragmatic inconsistency has profound implications. Drawing on speech-act theory as developed by Austin and Searle, Apel conceptualizes the respective kind of inconsistency as a "performative self-contradiction." This notion, Apel contends, is the "post-linguistic-turn equivalent to what Kant considered a violation of the 'self-consistency of reason' [*Selbsteinstimmigkeit der Vernunft*]." ³

University Press, 1998), 43–63; Karl-Otto Apel, *Paradigmen der Ersten Philosophie: Zur reflexiven—transzendentalpragmatischen—Rekonstruktion der Philosophiegeschichte* (Frankfurt: Suhrkamp, 2011).

² Cf. the two most important collections to date of Apel's essays: *Auseinandersetzungen in Erprobung des transzendentalpragmatischen Ansatzes* (Frankfurt: Suhrkamp, 1998), and *From a Transcendental-Semiotic Point of View*, ed. Marianna Papastephanou (Manchester, Manchester University Press, 1998). In addition to elaborations on the consensus theory of truth, post-metaphysical foundationalism, and ethical relativism, there is a momentous critique of Habermas's discourse theory of democracy. Apel's *Diskurs und Verantwortung: Das Problem des Übergangs zur postkonventionellen Moral* (Frankfurt: Suhrkamp, 1988) develops a non-Habermasian architectonic for the justification and application of discourse ethics; the long and as yet untranslated closing essay contains the gist of Apel's appropriation of Lawrence Kohlberg's cognitive-psychological account of moral development. *Selected Essays, Volume One: Towards a Transcendental Semiotics*, ed. Eduardo Mendicta (Atlantic Highlands, NJ: Humanities Press, 1994) contains Apel's objections to Davidson's truth-conditional semantics and Searle's intentionalist semantics; unfortunately, the translation is uneven.

³ Apel, "What is Philosophy?," 156.

Apel (unlike Habermas) elaborates from the notion of a performative self-contradiction a method of rationally definitive justification (*Letztbegründung*): testing propositions that are being asserted with a universal claim to validity for their pragmatic consistency in the very act of proposing them “provides the most radical criterion of sense-critique” (*Sinnkritik*) and is thus “the primordial method of a transcendental critique of reason after the linguistic turn of philosophy.”⁴ The method of rationally definitive justification is not only critical, or limiting, it is also, as Apel suggests,⁵ establishing, or positive, since we can use it in order to reveal presuppositions that are necessary for any validity-fixing argumentation.

Apel deserves to be better known as the originator of discourse ethics (*Diskursethik*), because he, not Habermas, had already mapped out the basics of this framework of meta-ethics and normative ethics by the late 1960s in a seminal essay about the constitutive role of the idea of an unlimited community of communication with regard to meaning and interpretation.⁶ The initial plausibility of discourse ethics depends on the intuition that the dialogical practice of fully engaged argumentative discourse necessarily involves conceptually normative presuppositions and that some of these have a recognizably moral content in the form of requirements for which universal validity can be claimed. The requirements concern discourse partners and are recognizably moral in the sense that they reciprocally expect that discourse participants ought to take seriously how the consequences of their argumentative acts affect all other reason-responsive agents for their good or ill.

⁴ Apel, “What is Philosophy?,” 167. Cf. “The Problem of Philosophical Foundations in Light of a Transcendental Pragmatics of Language,” in *After Philosophy: End or Transformation?*, ed. Kenneth Baynes, James Bohman, and Thomas McCarthy (Cambridge, MA: MIT Press, 1987), 250–90.

⁵ Apel, “What is Philosophy?,” 167.

⁶ See “The A Priori of the Communication Community and the Foundations of Ethics,” in *Toward a Transformation of Philosophy*, trans. Glyn Adey and David Fisby (London: Routledge, 1980), 225–300 (reprinted by Marquette University Press, 1998). Cf. Karl-Otto Apel, *Transformation der Philosophie*, vol. 1, *Sprachanalytik, Semiotik, Hermeneutik*, and especially vol. 2, *Das Apriori der Kommunikationsgemeinschaft* (Frankfurt: Suhrkamp, 1976).

Not surprisingly, many nontrivial problems have been noted in Apel's TPC approach generally, and also specifically in discourse ethics within this approach. The list of objections that have been raised to Apel's and Habermas's related versions of a philosophy of communicative pragmatism would be long indeed, and a metacritical survey of the many critical discussions that both approaches have already spurred I must bracket for the purpose of the present chapter. Instead, I want to review Apel's original intuition: the intuition that by reflexive recourse to our practices of argumentative discourse, we can ground in a rationally definitive way certain normative requirements (moral and other) as requirements that any and all agents, in their capacity of enacting communicative rationality, must meet and endorse.

Transcendental Reflection

The preceding brief introduction to the key ideas of TPC should have made it clear that Apel interprets transcendental reflection as an awareness of necessary elements within the actual self-understanding of participants in argumentative discourse. The concept of such an awareness is not a concept of psychology, like that of introspection. Rather, it is a concept of first philosophy, like rational necessity or truths of reason. For Kant, transcendental reflection determines "the origins of key cognitive representations in sensibility, understanding, or reason, and their a priori roles and relations in cognition, and thus their contributions to the possibility and validity of knowledge."⁷

How about transcendental arguments? Any argument that is suitable for providing a convincing route for transcendental reflection so conceived can be called, relative to the Kantian enterprise, a "transcendental" argument. The long and winding search for specifically transcendental arguments as a distinct form of arguments in anglophone analytic philosophy and in continental philosophy has not produced any philosophically

⁷ Kenneth R. Westphal, "Epistemic Reflection and Transcendental Proof," in *Strawson and Kant*, ed. Hans-Johann Glock (Oxford: Clarendon Press, 2003), 130.

encouraging results.⁸ Candidate skeletal forms that have been proposed apparently do not differ from the standard format of reasoning by modus ponens, except perhaps in the content of their premises.⁹ An altogether not implausible interpretation of these meager results is that the hope of locating the very point of transcendental reflection in a clearly distinct inferential format—transcendental arguments—is misguided. Maybe transcendental reflection has a specific point but no equally specific form of argument that would serve to establish that point.

For Apel, transcendental reflection is what discloses those normative commitments, roles, and relations in the reciprocally shared self-understanding of discourse participants that are indispensable for the potentially rational powers of argumentative discourse. Any argument that serves as a vehicle for transcendental reflection so conceived we can call, relative to the Apelian TPC enterprise, a “transcendental” argument. As pointed out in the introduction, Apelian transcendental arguments depend essentially on the notion of a performative self-contradiction or, put differently, they depend essentially on the principle that interlocutors in discourse must not permit performative self-contradictions. Interlocutors must in principle be performatively self-consistent.

In order to trace the connection between performative self-consistency and the aim of Apelian transcendental reflection (namely, disclosing indispensable normative commitments, roles, and relations in the reciprocally shared self-understanding of discourse participants), we first have to register the anti-Cartesian thrust of TCP. TCP attributes rational powers primarily to a dialogical *practice*, namely argumentative discourse, and not

⁸ Cf. Robert Stern, “Transcendental arguments: A Plea for Modesty,” *Grazer philosophische Studien* 74, no. 1 (2007): 143–61; reprinted in *Philosophical Knowledge: Its Possibility and Scope*, ed. Christian Beyer and Alex Burri (Amsterdam: Rodopi, 2007). Cf. also Marcel Niquet, *Transzendente Argumente: Kant, Strawson und die Aporetik der Detranszendentalisierung* (Frankfurt: Suhrkamp, 1991).

⁹ Chase and Reynolds, for instance, give the following characterization that allegedly captures what is distinctive of a transcendental argument: “(1) Subject-involving state of affairs p obtains. (2) A necessary condition for p obtaining is that q obtain. (C) So q obtains.” James Chase and Jack Reynolds, “The Fate of Transcendental Reasoning in Contemporary Philosophy,” in *Postanalytic and Metacontinental: Crossing Philosophical Divides*, ed. Jack Reynolds, James Chase, James Williams, and Ed Mares (London: Continuum, 2010), 29.

primarily to the *minds* of individual persons construed solipsistically as “Cartesian subjects,” as a powerful philosophical tradition would have it.

The rational powers of argumentative discourse, according to Apel, consist in its capacity to realize the potential intersubjective validity of any kind of thought content fit to be asserted with the communicative intention of claiming its validity vis-à-vis a community capable of communicating and reasoning. The rational power of real argumentative discourse consists in determining the relative validity of some determinate thought content that some determinate interlocutors in their argumentative roles of proponent and opponent can intend to posit, revise, or retract.

A determinate thought content that someone posits (for example, when interlocutor S truth-committedly asserts p: that such and such is the case) is intersubjectively valid only in relation to reasons, reasons that are good enough to back the corresponding validity claim (for example, S’s claim that p is true) if challenged by other interlocutors in discourse. Only if such reasons can be given is there a *reason-relative justification* for making and maintaining the respective validity claim and for complying with it in the absence of good defeating reasons. Validity claims that someone might raise but for which no reasons at all can be given are rationally indeterminate, and no one is rationally justified in making them or expecting compliance with them vis-à-vis anyone else.

Perhaps the terms “thought content” and “validity claim” that I choose in order to interpret Apel’s TPC at this point might ring slightly strange against the backdrop of mainstream talk of propositions and their truth conditions. However, it is essential to make explicit the idea in both Apel’s TPC and Habermas’s formal pragmatics that the phrase “validity claim,” as a translation of the German term *Geltungsanspruch*, “does not have the narrow logical sense (truth-preserving argument forms), but rather connotes a richer social idea—that a claim (statement) merits the addressee’s acceptance because it is justified or true in some sense, which can vary according to the sphere of validity and dialogical context.”¹⁰ Certainly,

¹⁰ As James Bohman and William Rehg put it in “Jürgen Habermas,” *The Stanford Encyclopedia of Philosophy* (Fall 2014 edition), ed. Edward N. Zalta, <http://plato.stanford.edu/archives/fall2014/entries/habermas/>.

propositions with truth conditions are paradigmatic for the kind of thought contents that speakers can posit via an established linguistic practice—the language game of fact-stating language—such that a universal claim to validity gets attached to them, turning a merely subjective opinion into a statement of facts whose claim to be either true or false is up for justification and criticism. But so are probably some thought contents whose core meaning is evaluative (for example, “that burning the cat on the mat was extremely cruel”) or normative (for example, “that cruelly inflicting pain on a sentient animal is morally wrong and impermissible”). Apparently, evaluative and normative thoughts too can be posited, that is, can be claimed to be convincing, owing to their reasons, in a generalizable way that commands assent intersubjectively within a more or less inclusive community capable of communicating and reasoning.

So far, I have recapitulated TPC’s doctrine of rational validity as emerging from communicatively constructed claims. Validity qualifies as rational validity by depending on reasons. The dependency relation, here, is a justificatory relation that permits basically two moves for participants in a discourse who scrutinize validity claims and their reasons: the affirmative move of endorsing a claim, and the critical move of rejecting a claim. The affirmative move results in a consensus that expresses reciprocally shared judgment; the critical move, if maintained, results in a dissent that expresses that reciprocally shared judgment is lacking as yet. Both moves, of course, qualify as rational only if there are good reasons, from the vantage points of the participants themselves, for moving either in one way or in the other way.

Let us note in passing that unfortunately Apel nowhere gives a detailed discourse-theoretical explanation of validity and its varieties.¹¹ Let us suppose for the sake of argument that the outline of such a theory that is found in TPC could be fully fleshed out. Then what work (if any) does Apelian transcendental reflection do in all this?

Consider again: according to TPC, the rational power of real argumentative discourse consists in determining the relative validity of some determinate thought content that some determinate interlocutors in their

¹¹ Nor does Habermas, after retracting his early inchoate “consensus theory of truth.”

argumentative roles of proponent and opponent can intend to posit, revise, or retract. If our practices of argumentative discourse really can have, or under realizable conditions would have, these powers for all persons that are capable of communication and reason like us, and if we want to know what is constitutive of this real possibility, then we want to know what is necessary for taking this to be a real possibility in our discourse.

Now add to this line of inquiry the condition that our starting point is from within, that is, from us here and now in our own proper instance of discourse. Call this the “performative condition.” Without this condition, we would pursue the cultural anthropologist’s business of reconstructing how practitioners make sense of their practices. With this condition, however, we are out to clarify how we ourselves actually make sense of our own practice. Call this “internal reconstruction.” This is not about how certain people conceive of a practice they call discourse. This is about how we understand discourse, and how we understand ourselves within discourse, when we are actually involved in it. Here, our cognitive access to our ongoing meaningful practice is as direct as it gets.

Note that once we claim validity for whatever insights we think we received by way of internal reconstruction, we have to invoke and enact the very practice whose composition we reconstruct. For how else could we claim to know anything, if not by arguing for the rational validity of our claims? Direct cognitive access to ongoing argumentative discourse is a reflexive mode—or, if you will, a self-referential mode—of cognitive access: whatever it is that we validity-claimingly come to think about argumentative discourse as such, we must by the same token be also able to think validity-claimingly about the actual discourse in which *we* articulate *here and now* what we have come to think about argumentative discourse as such.

That we “must by the same token be also able to think,” here, expresses generic requirements of consistency and coherence. That “we” must be able so to think expresses the fact—according to TPC it is a fact, albeit of reason—that validity is an intersubjective affair: I cannot address a validity claim to myself in particular; validity claims address others besides myself in the same sense in which I am their addressee. They must allow for a reciprocity of recognition from the

points of view of some other discourse participants and, in the case of “universal” validity claims (as in claiming that what someone asserts is true or false), from the points of view of conceivably *all* other discourse participants.

Apel’s aspiration that transcendental reflection furnishes us with a means of rationally definitive grounding (*Letztbegründung*) is the aspiration that by direct reflexive access to actual argumentative discourse we ascertain certain normative requirements (logical, moral, and other) as requirements that any and all agents who share communicative rationality must posit and acknowledge. The flipside of the rational necessity of the “must posit and acknowledge” is the rational impossibility of sharing the communicative rationality of discourse participants while negating and repudiating the respective normative requirements. Transcendental arguments in this vein are arguments by which we establish the rational necessity of certain normative requirements in discourse, by disclosing the rational impossibility of attacking them successfully in discourse. Such transcendental arguments establish that something is rationally necessary by establishing that a justificatory move towards it, no less than a skeptical move away from it, is impossible without it. Without it, the success of the justificatory move, no less than the success of the skeptical move, is not conceivably possible. That is why the skeptical move away from it is necessarily unsuccessful and can be seen to be so. Transcendental reflection serves to demonstrate the limits of skepticism: radical skepticism that would disclaim something that is necessary for radical skepticism to be possible would transgress a limit beyond which the skeptical move loses its rational power to doubt anything. The skeptical move cannot simultaneously (i) eschew the medium of discourse whose rational powers provide its critical thrust and (ii) satisfy the conditions that are necessary if any discourse (including skeptical discourse) is to have rational powers at all (for instance, the power of skeptical doubt).

If transcendental reflection delivers on its promises, then we have achieved no small thing. After all, we are looking for the constitutive elements of that dialogical practice wherein it is possible for all agents (that is, all agents who recognize one another as capable of enacting communicative rationality) to justify to each other all of their validity

claims on each other. Evidently, this practice has utmost practical importance for beings like us. And it is hard to think of what it would be like to live in a world in which it would have lost its importance.

The Contours of Argumentative Discourse

The preceding section has shown that, in TPC, transcendental reflection is reflection by an agent of communicative rationality on the activity of argumentation from within actual episodes of the very activity of argumentation. And since the activity of argumentation, being dialogical, involves more than one social role, it takes more than one rational agent for its full realization.¹² Structurally required are thinkers/speakers capable of first, second, and third-person thinking, at least one in the role of an *author* of a validity-claiming utterance or thought, one in the role of an *addressee* of that utterance or thought with the potential to reciprocate, and a third thinker/speaker in the role of a *reporter*, who can report to other thinkers/speakers what the author claimed vis-à-vis an addressee and what the latter claimed in response. These roles are usually, but need not be, taken on by distinct real persons. Any competent thinker will be able virtually to take on those roles in turn by him or herself. Evidently, the activity of argumentation requires suitable dialogical practices in which different rational agents can act, and are willing to act, as communicatively connected rational agents, and to do so while sharing the aim of cooperatively probing, contesting, and defending determinate validity claims.¹³

¹² Argumentation is not an activity that any radically socially isolated individual could learn. Of course, once socially learned, argumentation can be carried out *in foro interno*, i.e., embedded in the first-personal thought of a single person.

¹³ In terms of a distinction that Habermas has made popular, one could express this point by saying that rational agents, when involved in discourse, must be able and willing to act communicatively, not strategically or instrumentally. Cf Jürgen Habermas, "Discourse Ethics: Notes on a Program of Philosophical Justification," in *Moral Consciousness and Communicative Action*, trans. Christian Lenhardt and Shierry Weber Nicholsen (Cambridge, MA: MIT Press, 1990), 43–115; Jürgen Habermas, *Justification and Application: Remarks on Discourse Ethics*, trans. C. Cronin (Cambridge, MA: MIT Press, 1993); Maeve Cooke, *Language and Reason: A Study of Habermas's Pragmatics* (Cambridge, MA: MIT Press, 1997).

Apel's somewhat monotonous way of making this point about the interdependency of dialogical practices, validity, and communicative connectedness is by repeating that to identify with a "virtually unlimited community of communication" is "an *a priori*" requirement for rational agents in order for them to be able to think thought contents validity-claimingly. Apel's point is not merely a psychological one about identification, nor is it merely a sociological one about practice. Rather, the point about the *a priori* status of situating ourselves at once in both our finite real community of communication and in a virtually unlimited community of communication is a conceptual point about the concept of rational validity: if it *is* rationally valid to claim *that p* (a fact-expressing thought) *is true*, or *that n* (an obligation-expressing thought) *is right*, then the rational validity of so claiming must outstrip the particular set of particular individuals to whom this appears to be so. When we think, say, that it is rationally valid to claim that the earth is flat is true, or that forbidding the killing of innocent people is right, then it would be odd to think in the same regard that this is so only for us (and others) to whom these claims appear to be rationally valid claims. Rather, it would be natural to think that these claims appear to us (and others) to be rationally valid claims *in virtue of certain reasons* that sufficiently justify the respective claims and make anyone's claiming validity for these reasons into rationally valid claimings. Rational validity extends as far as the reasons can be shared on which such claiming is based. Likewise for disclaiming that a certain claim that purports to be rationally valid really is so.

The very idea of claiming validity (be it affirmatively or critically) is the idea of a *social* practice. And to think anything validity-claimingly is to include this idea of a social practice of the right kind internally into the thought content one is thinking. But what makes a social practice of dialogical exchange a social practice of the right kind? How do we know? Who is to judge? There does not appear to be any other direction from which to expect an answer to these questions than the direction that a constitutive aim of argumentative discourse would provide. But does discourse have a constitutive aim? Whenever we are fully engaged in discourse, do we have a constitutive aim by being so engaged? Once again: How do we know? And who is to judge?

A nonconstitutive aim of a type of practice would be an aim that, if dropped, would leave the practice in its proper function. A constitutive aim would be an aim that could not be dropped in the practice while leaving its spirit unscathed.¹⁴ As a matter of fact, we can use practices of argumentation for multifarious purposes, for instance for fighting. However, to drop the aim of reaching a rational (that is, reason-responsive) consensus amongst discourse participants about the claims whose rational (that is, reason-responsive) validity they probe, it seems, would be to lose entirely what practices of argumentation should be good for. It appears safe to claim that this is the constitutive aim of argumentative discourse. And if someone else were to claim that this is not so, then in order to take this different opinion seriously, we would have to embark on a practice that already has at its aim the reaching of a rational consensus about conflicting claims amongst discourse participants. And there appears to be no practical equivalent of argumentation as we know it that could be trusted to lead us towards this aim. Rational validity claims cannot be fixed, apparently, other than by way of argumentation.¹⁵ Of course, it is not certain that in every case we can and will actually reach the constitutive aim of argumentative discourse: a determinate consensus on a determinate validity claim. The constitutive aim of argumentative discourse is a (not altogether unrealizable) practical ideal.¹⁶

Since argumentative discourse is not a type of speech act (like, for example, the speech act of assertion) but a rich practice for which many kinds of speech acts have to be available besides those of assertion, questioning, answering, stating, doubting, and so on, discourse cannot be contrasted to nonargumentative dialogical practices in terms of norms

¹⁴ Whereas the distinction between nonconstitutive and constitutive rules concerns fixing the identity of practices (e.g., fixing the identity of the game of chess by fixing the rules of chess), I want the distinction between nonconstitutive and constitutive aims to concern fixing the normative essence of practices, that is, fixing how a practice must aspire to be in order to serve its aim well. Drop the aim and you can no longer know how the practice aspires to be.

¹⁵ This is not to say that there are altogether no other ways for fixing our validity beliefs, for example by threats or cognitive conditioning. But these other ways can be faulted for irrationality when our question is to find out what *is* valid and what *is not*.

¹⁶ How this relates to the way Kant distinguishes between “regulative” and “constitutive” ideas cannot be discussed within the confines of the present chapter.

alone, or effects alone, or commitments alone.¹⁷ Considering that argumentative discourse is a practice above the level of speech acts, we may expect norms, effects, commitments, and perhaps further determiners to contribute jointly to giving argumentative discourse its characteristic shape. But essentially it is the *aim* of discourse which gives discourse its *proper* shape. Argumentative discourse, as Apel uses the term, is a dialogical practice for exchanging arguments, in which all participants enact communicative rationality and are oriented towards reaching a consensus about contested validity claims concerning a determinate thought content that is, in its linguistic articulation in the form of statements, the referent of their discourse.

Apel has less to say about how the fixing of validity claims by argumentatively intending a rationally prompted consensus about them actually operates. I propose to augment Apel's characterization of the essence of argumentative discourse, as distinct from nonargumentative discourse or nondiscursive dialogical practices. I propose to conceive of discourse as the practice in which revision (or re-evaluation) of conflicting reasons is permanently possible and must remain possible in order for us to engage in this very activity. Furthermore, the revision (or re-evaluation) of conflicting reasons proceeds by bringing to bear further reasons: the operation of revising or re-evaluating conflicting reasons has to be enclosed within the space of reasons if this operation is to be governed by procedures which the persons who are following them want to be such that they cannot be faulted for irrationality. But changing my evaluation of any reasons as a consequence of, say, chance, threats, forgetfulness, social mimesis, distraction, or neglect, would not count as revising them. In order to revise reasons in which I have rational stakes, I have to change my evaluation of them guided by nothing else than my unfettered insight, that is, in light of other of my reasons whose rational credentials I take, for the time being, as unquestionable or at least as less questionable than the reason that has come under scrutiny for revision.

¹⁷ Brown and Cappelen survey various attempts to capture the normative nature of assertions in these terms, cf. "Assertion: An Introduction and Overview," in *Assertion: New Philosophical Essays*, ed. C. Brown and H. Cappelen (Oxford: Oxford University Press, 2011), 1–17.

It is in this sense that I would like to characterize argumentative discourse essentially in the following way:

(D) *Argumentative discourse is a social practice, open to all persons in their capacity as reasonable evaluators, which has as its aim the communicatively rational revision of conflicting reasons with (apparently less conflicting) reasons.*

Note that in shifting from reasons considered specifically as justifiers of the three or four “universal validity claims” to which Apel and Habermas repeatedly refer (truth, rightness, veracity, and meaning) to reasons as such, I have arrived at a more general characterization of argumentative discourse.

This has more than one advantage. One is that we can bracket and reserve for further clarification two complex interrelated problems. The first problem that remains unsolved in TPC and in its Habermasian counterpart is the problem of providing a robust taxonomy of validity claims. The second problem not yet satisfactorily settled is the problem of the scope of validity claims. Clearly, the paradigm of a validity claim of unrestricted scope is the claim that a fact-stating proposition is true, or false. Consider: a truth claim is unrestricted in the sense that to claim that *p* is true is to claim that what *p* states as factually being so really is so, no matter whether only some (we here and now) or all others too grasp the corresponding fact. In TPC, this point is expressed in terms of universality: if a determinate truth claim is valid, its truth (that is, the truth of its content) may in principle be claimed vis-à-vis any and everyone, that is, *universally*. The qualification “in principle” must be added here, since it is obviously not the case that every person is able to grasp every fact that every other person can grasp; nor are all situations, in which truth is a relevant consideration per se, situations in which truth must be claimed and falsity criticized, even when this would be possible for someone involved in the respective situation. Now, Habermas wavers about whether veracity is a universal or a non-universal validity claim. He is committed to taking moral rightness as a universal validity claim that is, as both Habermas and Apel

suppose, truth-like.¹⁸ However, neither Apel nor Habermas provide more substantive discussions of their debatable position, over and above appeals to examples that give it some plausibility. Technically speaking, TPC and its Habermasian counterpart strive for a position of alethic pluralism in the theory of truth—that there is more than one truth property across different domains of discourse—and both are inflationists, in that both hold that the concept of truth does not boil down to what is captured in a truth-conditional analysis in the Tarskian vein. But much remains to be clarified.¹⁹

Another major advantage of my more general characterization of argumentative discourse is that it suggests a clear operative role for discourse participants. For rational agents to act as argumentative discourse participants is to act as reasonable evaluators of reasons, which they do in light of other reasons. This serves to bring out a continuity of acting for a reason in all kinds of situations and acting in discourse. Suppose I think that I have reason R for doing A in situation C, and I think that my doing A in C is reasonable in virtue of my doing it for reason R (and not for some other reason). Suppose someone criticizes not only what I do but how I understand what I do. Let us assume that we act in suitable social conditions that support the move to argumentative discourse in the face of conflicting convictions. In that case, questioning my taking R to be a good reason for doing what I do will lead us to take up argumentative discourse, such that we now try to establish R's *true value* by evaluating R, my supposedly good-enough reason for acting

¹⁸ For an overview of the development of Habermas's position, see Maeve Cooke's Introduction (1–19) in Jürgen Habermas, *On the Pragmatics of Communication*, ed. Maeve Cooke (Cambridge, MA: MIT Press, 1998).

¹⁹ For a range of issues that are at stake in the inflationist versus deflationist debate, cf. Corey D. Wright and Nikolaj J. L. Pedersen, eds, *New Waves in Truth* (New York: Palgrave Macmillan, 2010). Apel gives the most detailed statement of his truth-theoretical views in an essay in which he attempts to integrate correspondence, evidence, coherence, and consensus into a comprehensive truth property; see "Fallibilismus, Konsensstheorie der Wahrheit und Letztbegründung," in *Auseinandersetzungen in Erprobung*, 81–193. Cf. "Transcendental Semiotics and Truth: The Relevance of a Peircean Consensus Theory of Truth in the Present Debate about Truth Theories," in *From a Transcendental-Semiotic Point of View*, 64–80.

in C, by the measure of other reasons R^* that are serviceable as acknowledged standards of evaluation in our case. For instance, I might justify and defend my action by disclosing that R appears to me to be a perfectly good reason in light of standards of morally permitted prudence. You might object that moral considerations are irrelevant for A in C, and that when we consider and evaluate R for doing A in C by standard reasons of pure prudence, R will not count for much. Reasoning about arguably good reasons with a view to consensually settling their true value is already reasoning in the mode of argumentative discourse. Reasoning about the relevance and ranking of determinate standard reasons that we want to bring to bear on targeted first-order reasons is part and parcel of what reasonable evaluators of reasons do when engaged in argumentative discourse. Claims concerning the relative rational merits or demerits of purportedly “good reasons” are the most elementary format of rationally justifiable validity claims.

Performative Self-contradictions and the Transcendental Kernel of Discourse Ethics

With an augmented concept of the practice of argumentative discourse in place, we can now turn to the question, very prominent in Apelian TPC, whether some of D’s multifarious prerequisites (1) are conditions of its very possibility such that they are necessary for its constitutive aim, and (2) such that their being necessary for its constitutive aim can be brought out in transcendental reflection. Intimately connected with this question are the chances that a discourse ethics can get off the ground. As was stated in the introduction, Apelian discourse ethics hinges on the intuition that the dialogical practice of fully engaged argumentative discourse necessarily involves conceptually normative presuppositions, and that some of these have a recognizably moral content in the form of requirements for which universal validity can be claimed. How can one go about defending this intuition as an insight?

The identification of morally charged presuppositions of discourse and the proof that Apel offers for their unassailable rational credentials both rely crucially on arguments of a transcendental bent:

For any reasonable evaluators P1, P2, P3, as author, addressee, and reporter in an ongoing practice of discourse, if abandoning a certain concept *c* would clash with the very possibility for all of them of construing their common practice of discourse as rationally meaningful, then *c* is conceptually necessary in the sense of requiring use in all possible worlds in which discourse as we know it exists. In that sense, all the concepts that we have to operate with in order to understand ourselves as entokening the essence of discourse are such conceptual presuppositions of discourse.

If among *c* there is a conceptually necessary element of a *normative* kind (cn), permitting or requiring something of someone, and if for any reasonable evaluators P1, P2, P3, as author, addressee, and reporter in an ongoing practice of discourse the intention to disclaim that cn is valid would make it impossible for all of them to construe their common practice of discourse as collectively aiming at a consensus about contested reason-based validity claims about the reference of their discourse, then to acknowledge *n*'s validity is necessary in all possible worlds in which discourse as we know it exists.

At this point we might want to insist on a principal difference between *our acknowledging* cn's validity and cn's having validity or *being* valid. However, recall TPC's doctrine of rational validity, which emerges from communicatively constructed claims and is nothing over and above a construction of communicatively related rational agents. If we accept this doctrine of rational validity, we see that no real difference between acknowledging cn's validity and cn's validity can be made when its validity could not be doubted without by the same token acknowledging its validity (since doubting anything requires discourse, and discourse requires acknowledging that cn is valid).

Consider D, the social practice, open to all persons in their capacity as reasonable evaluators, which has as its aim the communicatively rational revision of conflicting reasons with (apparently less

conflicting) reasons. What general assumptions, or lack of general assumptions, would unhinge our self-understanding as entokening the essence of discourse?

For a start, we have to attribute to other persons no less of the generic rational capacities that we think we have and that it takes in order to engage in D at all.²⁰ Evidently, a sufficient command of certain linguistic and cognitive competences is a conceptual presupposition of D, since lacking the respective competences we could not gain or maintain a self-understanding in our practice as entokening the essence of discourse. Prominent among the elementary logic that participants in discourse must reciprocally presuppose to take seriously will be the avoidance of open logical contradictions. For if open logical contradictions were nothing to be avoided, then no reasons would ever stand in need of revision.

If we knocked out as a presupposition the rule to take potential others in their capacity as evaluators of reasons neither more nor less seriously than we actually take ourselves, then we could no longer think of ourselves as aiming at communicatively rational revision of conflicting reasons, since conflict would disappear. Only for reasoners on an equal footing can their reasons conflict. And only for reasoners on an equal footing can it make sense to aim at fixing the true value of their reasons by way of D.

Apel favors an articulation of the moral relevance of some of the necessary presuppositions of argumentative discourse in terms of a moral co-responsibility (between all actual as well as all possible participants in discourse) for keeping all their discourse-pertinent actions in accordance with the generic deontic status of free and equal co-subjects.²¹ The generic deontic status of free and equal co-subjects ideally requires of whomever is involved in argumentation the normative attitude of wanting all interlocutors to accept a default stance of

²⁰ In specialized discourses, such as expert discourses, we might want to exclude some reasonable evaluators and rest with the set of those whom we treat as our discursive peers concerning the issue at hand. But the selective exclusion is rational if based on a reason that can be offered to them, and accepted by them, as a justification of their exclusion.

²¹ See Apel, *Diskurs und Verantwortung*.

mutual recognition and symmetrical situatedness as binding on anyone who is or could be so involved. Apel holds that it is at least implicitly recognized in any serious argumentative discourse that discourse (be it a purely theoretical discourse about truth claims or any other kind of validity claims) presupposes a minimum of moral rules, a “discourse ethics”: “the rules of the game or norms of an ideal discourse community.”

In TPC, what counts as the most convincing way of demonstrating the validity of the normative content of conceptually necessary presuppositions of D is by demonstrating that whoever attempts to deny that they are valid cannot validly do so, since doubt that could make a difference in discourse requires that their validity is acknowledged.

How can we demonstrate that this is so? It is at this point that the notion of a performative self-contradiction becomes crucial. Such a self-contradiction does not reduce to an inconsistency that holds between two potentially truth-bearing contents, as in a conjunction of logically contradictory propositions. A performative self-contradiction is a logical-cum-practical inconsistency between, on the one hand, a determinate content *c* that some speaker *S* posits with a claim to universal validity, and on the other hand, at least one of those of *D*’s normative presuppositions that pertain to *D*’s constitutive aim.

In other words, the notion of a performative self-contradiction is the notion of a predicament that any rational evaluator like us would basically want to avoid in argumentation. We attribute a performative self-contradiction to a discourse participant *P1* if *P1* intends as author to claim universal validity for content *c*, and *c* is such that if it *were* universally valid then we (*P2* and *P3* as addressee and reporter) could not sensibly attribute to *P1* the very intention to claim universal validity for *c*.

Let content *c* be the thought that one is permitted to assert whatever one wants if that furthers one’s interests. If *P1* intends to claim validity for a corresponding statement, we cannot take *P1*’s speech act of asserting *c* seriously (as we would have to in order to understand *P1* and ourselves as discourse participants), since if *P1* were right (which we took him to aim at) then we all would be allowed to assert whatever we want

and convincing someone that someone is right (or wrong) would no longer be, and would never have been, a sensible intention to form.

Let content *c* be the thought that judging reasons for their merits and demerits is a subjective affair and totally arbitrary. P1 purports to articulate this thought in the form of a statement addressed to us: "Judging reasons for their merits and demerits is a subjective affair and totally arbitrary." Taking P1 seriously in what he says, we must understand him as aiming at stating something true. Yet if what he seems to state *were* really true, then whatever reason he could have for being himself convinced and for convincing us will have no true value at all and hence cannot make any difference in discourse.

Let me (P1) address you (P2) with the following eliminatively reductionist belief: "Argumentation and this entire game of giving and taking reasons is nothing but a power game for winning others over to one's own opinions." You (P2) report to her (P3): "P1 claims that argumentation is nothing but a power game." It would be natural for P3 to respond skeptically: "Why should this be true? Why does P1 think *that*?" But we all would know that no reason that could be convincing could possibly be forthcoming if what P1 thinks and says were really true, and that no reason ever could have been.

To see why a performative self-contradiction is neither merely a logical contradiction nor merely a practical infelicity but instead a logical-cum-practical inconsistency, think reflectively about what it is like for me actually to contradict myself performatively. Between us as potential discourse partners, I present myself to you as willing *to make an appropriately reason-grounded validity claim concerning some judgeable content c* such that you and I should accept this claim as insightfully warranted. But the posited content *c* is such that if what I claim were valid in the way I seem to intend (for example: *true, morally right, legitimate in some other-than-moral sense*) then that very intention (namely, *to make an appropriately reason-grounded validity claim concerning c*) would not be available to me as my intention. So now you have to attribute to me a logically inconsistent pair of beliefs: that I do believe *and* that I do not believe that I intend *to make an appropriately reason-grounded validity claim concerning c*. My practical move in the practice of discourse creates reasons that cancel my position as a participant in the practice of discourse.

In Conclusion: The Morality in Discourse

The TPC principle that discourse participants must avoid performatively contradicting themselves, and the logical principle that we must avoid believing open logical contradictions, are equally basic and complement each other in the constitutive aim of discourse.

Is the principle that performative self-contradictions must be avoided a valid principle of reason? Certainly yes, if we understand that reason involves communicative rationality. To intend to permit performative self-contradictions is impossible when at the same time you intend uptake as a discourse participant, i.e. being acknowledged as a person who has reasons that might convince us of something.

What is bad about actually contradicting oneself performatively? Or: What is bad about actually contradicting oneself logically? In both cases, willingly going against the relevant principle will cancel your status as an interlocutor whom we should respect as a co-equal within argumentation. In other words: you count yourself out as a reasonable evaluator of reasons.

So what is really bad about permitting contradicting oneself performatively as legitimate?

Or: What is really bad about permitting contradicting oneself logically as legitimate? A plausible answer takes the form of specifying a loss. Something important would break down, namely, argumentation—as a cooperative enterprise reaching a rational (that is, reason-responsive) consensus amongst discourse participants about the claims whose rational (that is, reason-responsive) validity, in principal, all discourse participants can probe, revise, and trust. What we stand to lose if argumentation mutated into schmargumentation²² is the real possibility of continually improving the stock of reasons that can serve as common grounds in all of our judgmental practices. One might be willing to accept this loss. But one knows that one could not will to accept this loss for any good reason.

²² Enoch raises serious doubts as to the prospects of deriving normativity from constitutive elements of rational agency; see David Enoch, “Agency, Shmagency: Why Normativity Won’t Come from What Is Constitutive of Action,” *Philosophical Review* 115, no. 2 (2006): 169–98. I think that Enoch’s criticism does not apply to the attempt of situating an important part of normativity in the constitutive aim of argumentative discourse.

I conclude with some remarks on the prospects of discourse ethics.²³ We have seen that those normative requirements that are part of the constitutive aim of discourse that are recognizably moral—the moral in discourse—require a default stance of equal and mutual recognition as a person capable of evaluating reasons reasonably. And these normative requirements are as secure as the practice of discourse itself. However, the morality intrinsic in discourse (MID) is too thin a normative moral system to be on a par with full-blown moralities as we know them, such as the normative moral system of “common sense morality.”²⁴ Confronting a diversity of moralities, the moral point of view that is associated with MID provides reasons for discriminating between those moralities that contain moral prohibitions against subjecting their constitutive beliefs to argumentative scrutiny, and those moralities that permit their constitutive beliefs to be examined in discourse. Call the latter *discourse-friendly moralities* and the former *discourse-averse moralities*. If the spirit of some particular morality M invites critical scrutiny of its constitutive beliefs in discourse and survives unscathed, we would appreciate M as a *discursively robust morality*.

MID provides good reasons for preferring discourse-friendly moralities to discourse-averse ones. From within MID’s point of view, one has good reasons to judge that a discourse-friendly morality is *morally* better than a discourse-averse one. However, the comparative “morally better,” here, projects these reasons only to one’s moral peers as defined by MID itself, that is, to persons in their capacity as reasonable evaluators of reasons: as valid discourse participants.

²³ I have elaborated my revisionist account of discourse ethics in Matthias Kettner, “Gert’s Moral Theory and Discourse Ethics,” in *Rationality, Rules, and Ideals: Critical Essays on Bernard Gert’s Moral Theory*, ed. Walter Sinnott-Armstrong and Robert Audi (Lanham, MD: Rowman & Littlefield, 2003), 31–50; and Matthias Kettner, “Discourse Ethics: Apel, Habermas, and Beyond,” in *Bioethics in Cultural Contexts: Reflections on Methods and Finitude*, ed. Christoph Rehmann-Sutter, Marcus Düwell, and Dietmar Mieth (Berlin: Springer, 2006), 299–318.

²⁴ Bernard Gert, *Morality: Its Nature and Justification* (Oxford: Oxford University Press, 1998).

We can of course cherish MID as the focus of a thin but maximally person-inclusive rational ethos.²⁵ This rational ethos, the ethically augmented morality in discourse, may properly be denoted by the term “discourse ethos” or “ethos of discourse.” The *ethos of discourse* is a morally normative stance, a particular conception of the moral. It is the moral stance that all committed proponents of discourse ethics will find congenial and recommendable. Such proponents make the moral integrity of the powers of discourse their foremost moral concern, and they extend this concern to extant formats and fora of actual communities of argumentation in our actual world.²⁶

According to Apel—and here we are back to the strong program of discourse ethics within TPC—the only unassailable hope for true moral progress is hope in the progressive globalization of the ethos of discourse.²⁷

Matthias Kettner is Professor of Philosophy at the University of Witten/Herdecke. He specializes in transcendental pragmatics, philosophy of culture, applied ethics, and philosophy of psychoanalysis. He is the author of *Reflexionen über das Unbewusste* (Vandenhoeck & Ruprecht, 2010) and has edited *Wunscherfüllende Medizin: Ärztliche Behandlung im Dienst von Selbstverwirklichung und Lebensplanung* (Campus Verlag, 2009) and *Angewandte Ethik als Politikum* (Suhrkamp Verlag, 2000).

²⁵ If it is true to claim that everyone must cherish MID as their rational ethos cannot be denied *salva rationalitate* (that is, without getting involved in a performative self-contradiction), the same is *not* true for the claim that one’s rational ethos must dominate all one’s ethical convictions. Hence, identifying with the discourse ethos, if it is required at all, it cannot be required by communicative reason *alone*.

²⁶ For instance, to ethics committees. For a discourse ethics approach to clinical ethics committees see Matthias Kettner, “Ein diskursethisches Beratungsmodell für klinische Ethik-Komitees,” in *Ethikkonsultation heute—Vom Modell zur Praxis*, ed. Ralf Stutzki, Kathrin Ohnsorge, and Stella Reiter-Theil (Berlin: LIT-Verlag, 2011), 45–57.

²⁷ Cf. Karl-Otto Apel, *The Response of Discourse Ethics to the Moral Challenge of the Human Situation as Such and Especially Today: Mercier Lectures* (Leuven: Peeters Publishers, 2001).

11

Transcendental Arguments Based on Question–Answer Contradictions

Yukio Irie

Language may be based on *relations of mutual responsiveness* among humans. What are the fundamental conditions for relations of linguistic mutual responsiveness? I would like to propose a method for addressing these problems by relying on a particular linguistic phenomenon that appears to reveal some fundamental conditions underlying linguistic communication. This phenomenon involves an odd kind of contradiction between a question and an answer, which can be illustrated as follows: “Can you hear me?”—“No, I cannot hear you.” I will call this type of contradiction a “question–answer contradiction” (QA contradiction). A QA contradiction seems to reveal fundamental structures or conditions of communication, because linguistic communication is essentially constructed out of question–answer relations; hence, the contradiction points to some basic or transcendental conditions of such relations. Through an

Y. Irie (✉)

Department of Philosophy, Osaka University, Osaka, Japan
e-mail: irie@let.osaka-u.ac.jp

© The Author(s) 2016

H. Kim, S. Hoeltzel (eds.), *Transcendental Inquiry*,

DOI 10.1007/978-3-319-40715-9_11

233

analysis of QA contradictions, I will discuss the normativity of language and the principles of identity, mutual belief, mutual recognition, and so on. First, I will classify and briefly analyze various types of QA contradictions. Second, I will propose a general analysis of question–answer relations (QA relations) in terms of anaphora and the presuppositions of questions, before applying this to the analysis of a QA contradiction. Third, I will use QA contradictions to make explicit some transcendental conditions for relations of linguistic mutual responsiveness.

Three Contradiction Types and the QA Contradiction

Conversation is essentially composed of questions and answers. Occasionally, however, a strange contradiction between questions and answers can arise. Although this contradiction may be rather familiar to readers, it has yet to be analyzed in detail.

We know of three types of contradictions: logical, semantic, and pragmatic. If a sentence or a set of sentences is *logically* contradictory, this means that it is logically impossible for the sentence or all sentences of the set to be true. In the case of a single sentence, the contradictory sentence usually has a form like “P and not P.” If a sentence or a set of sentences is *semantically* contradictory, this means that it is semantically impossible for the sentence or all sentences of the set to be true. A semantic contradiction is a contradiction arising from the meanings of words; an example of a single-sentence semantic contradiction is “This red rose is yellow.” A *pragmatic* contradiction is a contradiction between an utterance act (a phonetic act, a phatic act, and a rhetic act)¹ and its propositional content, or between an illocutionary act and its propositional content. Here is an example of a contradiction between an *utterance act* and its propositional content: “BE QUIET HERE PLEASE!” (uttered in a loud voice). Here are five instances of contradiction between an *illocutionary act* and its propositional content: (i) “I don’t exist,” (ii) “Don’t follow my order!” (iii) “I don’t make promises

¹ Cf. J. L. Austin, *How to Do Things with Words* (Oxford: Oxford University Press, 1962), ch. 8.

anymore,” (iv) “Words can’t express my gratitude!” and (v) “I declare nothing!” (Every type of illocutionary act could have a pragmatic contradiction.) A pragmatic contradiction involves two different domains, that of semantic meaning on the one hand and that of pragmatic action on the other. How do these two domains clash? We can see this clearly if we express the *utterance* or *illocutionary* act as a proposition; this proposition and that expressed by the uttered sentence contradict each other.

In the case of logical or semantic contradiction, if two sentences are contradictory to each other, then one is true and the other is false, but which one of these is false cannot be decided logically. Similarly, in the case of pragmatic contradiction between an *utterance* act and its propositional content, we have two ways to solve the contradiction: we can change the utterance act or its propositional content. However, in the case of pragmatic contradiction between an *illocutionary* act and its propositional content, what is falsified is always the propositional content of the uttered sentence, never the illocutionary act, because for the propositional content to be present, it must be uttered. Hence, the propositional content presupposes the utterance act of the proposition. Therefore, if these are in contradiction, then the propositional content must change to avoid the contradiction.

Pragmatic contradiction does not involve a contradiction at the level of a sentence or a proposition; rather, it involves a contradiction at the level of its use, that is, its utterance. Consider the following utterance: X says “I don’t exist.” Here, “I” refers to the utterer X. Hence the content of that utterance is the same as the content of the following utterance: X says “X doesn’t exist.” The former sentence is necessarily false, because it is false in every context in which it is uttered. But the latter sentence is contingently false, because its truth value depends on a particular context. In this context, that is, when uttered by X, it results in a pragmatic contradiction, because the propositional content is in contradiction with the illocutionary act. Whether an utterance constitutes a pragmatic contradiction depends crucially on the context of the utterance.²

²The message “I don’t exist” is true, for example, when one leaves this message for others to read or hear after one’s own death. Evaluated under these circumstances, the argument becomes more complicated.

In addition to these three types of contradictions, another kind can be found in certain relationships between questions and answers, such as: “Can you hear me?”—“No, I cannot hear you.” This negative answer cannot be characterized as a semantic or a pragmatic contradiction, but it is contradictory in relation to the question. This is what I call a QA contradiction.³ I will now describe two subtypes of QA contradiction with the use of examples.

Pure QA Contradiction

(1) *QA contradiction regarding the confirmation of physical connection.* The combination of a question and an answer, in which the latter confirms that the addressee cannot hear the speaker’s voice, seems to result in a QA contradiction. Consider again this example: “Can you hear me?”—“No, I cannot hear you.” The utterance “I cannot hear you” is not contradictory in itself. However, responding to the question with “I cannot hear you” is contradictory, because this answer would not have been offered if it were true. The question “Can you hear me?” cannot truthfully be answered by “I cannot hear you.” Therefore, if I answer truthfully, my answer must always be “yes.” The absence of an answer indicates that I could not hear the question or that I do not want to reply to it. (If the questioner thinks the addressee can hear him or her and nevertheless receives no answer, the questioner assumes that the addressee is refusing to answer. For this reason, individuals should reply immediately to prevent such a misunderstanding in typical cases.)

³ I initially identified QA contradictions as “discursive contradictions” in my paper “Paradox in Meta-communication (1)” [in Japanese], in *Bulletin of Osaka Shoin Women’s College*, 30 (1992). My second exposition on this subject was my paper “Some Contradictions between Questions and Answers” [in Japanese], in *The Ontology of Communication* (A Report of the Collaborative Research based on KAKENHI of JSPS submitted to JSPS in 2001, unpublished. Project Number 10410004). The third was my presentation “Contradiction in the Question–Answer Relation” at the 13th International Congress of Logic, Methodology and Philosophy of Science, Beijing, 2007. The fourth was my paper “Question–Answer Contradiction,” *Philosophia Osaka* 5 (2010): 79–87, where I distinguished between “the pure type of Question–Answer Contradiction” and “the ambiguous type of Question–Answer Contradiction,” which correspond respectively to the pure QA contradiction and the mixed QA contradiction of the current chapter.

In this case, the contradiction seems to be caused by the question, because the answer is not contradictory in itself. The question appears to involve pragmatic contradiction, because asking a question usually presupposes that an addressee can hear it, otherwise the question is not asked. It seems to be contradictory to inquire about such a presupposition. The illocutionary act of this question requires an answer about something that is typically presupposed as a condition of being able to address an addressee. Therefore this question seems to constitute a kind of contradiction: (i) the questioner doubts whether the addressee can hear the question, yet (ii) the questioner nevertheless attempts to address the addressee. These two attitudes seem to be contradictory, but not decisively contradictory. The relation may be deemed *nonconforming* in nature, rather than contradictory.

The negation of the proposition usually resolves the pragmatic contradiction. For example, the utterance “I don’t exist” represents a pragmatic contradiction, but its negation “I exist” does not constitute such a contradiction. In contrast, the *nonconformity* of questions cannot be resolved, even if their propositional content is negated. For example: “Can’t you hear me?”—“No, I cannot hear you.” This question is still nonconforming and the question and the answer still form a contradiction.

The contradiction between the question and the answer is determined not only by the nonconformity of the question but also by the negative answer. The propositional content of the negative answer contradicts the fact that an addressee is responding to the question.

One might wonder what the questioner should do when an addressee cannot hear the question and the illocutionary act of questioning does not succeed. In such circumstances, the questioner might not be intending to engage in a monologue. Indeed, dialogue may have been the goal. The speaker certainly intended to perform an illocutionary act, and this act failed. This case appears to be similar to the following situation. An individual throws a dart at a target, attempting, albeit without confidence, to hit it, and fails to do so. But these cases are quite different. Even though this questioner’s illocutionary act of questioning has failed (that is, no answer was forthcoming), she has succeeded in obtaining certain desired information (namely, that an addressee could not hear her).

(2) *QA contradictions regarding understanding a language.* If I am asked, for example, “Do you understand English?” and I answer, “No, I don’t,” this answer (not in the sense of “No, I don’t understand English well,” but in the literal sense) is contradictory, because the answerer must understand English in order to say “No, I don’t.” The utterance “I don’t understand English” is a pragmatic contradiction not only when it is uttered as an answer to the above question but also when it is uttered alone. However, answering that English question with “*Nein, ich kann Englisch nicht verstehen*” involves no pragmatic contradiction but instead a QA contradiction, because the responder must have understood the English question to provide this answer. This question must be posed in English for the question–answer pair to constitute a QA contradiction. For example, the following question and answer do not constitute a QA contradiction, even though the answer does constitute a pragmatic contradiction. “*Kannst du Englisch verstehen?*”—“No, I don’t understand English.”

Like the previous QA contradiction subtype, this case involves a *nonconforming* question. Generally speaking, a question is asked under the assumption that an addressee can understand the language used to ask the question. But the questioner doubts the validity of this assumption and seeks to confirm it with the question, articulating the following attitudes: (i) the questioner doubts whether the addressee can understand the language used to ask the question, but (ii) the questioner nevertheless addresses the addressee in that particular language. These two attitudes may not be contradictory, but they are nonconforming. The contradiction between the question and the answer arises not only from the nonconformity of the question but also from the negative answer. The propositional content of the negative answer contradicts the fact that an addressee is responding to the question uttered in that language.

(3) *Definition of the pure QA contradiction.* A QA contradiction involves a contradiction between a question and an answer to that question and, more precisely, between the propositional content of the answer and the responding relation of the answer to the nonconforming

question.⁴ I would like to classify QA contradictions of the sort exemplified above as *pure* QA contradictions. These are QA contradictions in which the answer does not involve pragmatic contradiction or, if it does, the pragmatic contradiction is not necessary and can be eliminated by a different rephrasing. The following examples are also cases of pure QA contradictions.

“Can you remember my question?”—“No, I can’t remember your question.”

“Can you answer my question?”—“No, I can’t answer your question.”

“Are you listening to me?”—“No, I am not listening to you.”

“How do you answer this question?”—“I don’t answer in any way.”

“Which language can you listen to?”—“I cannot listen to any language.”

Mixed QA Contradiction

(1) *QA contradictions regarding the confirmation of sincerity.* Engagement in a conversation involves a mutual understanding of mutual sincerity, which is constantly reconfirmed at the level of the metamessage. We can express this confirmation in the form of questions and answers as follows. “Are you talking to me sincerely?”—“Yes, I am talking to you sincerely”; or alternatively, “No, I’m not talking to you sincerely.” If the addressee intends to answer sincerely, he or she will answer with “Yes, I am talking to you sincerely.” However, even if the addressee does not intend to answer sincerely, he or she will still answer with “Yes, I am talking to you sincerely,” because otherwise the answer will in reality be true and sincere. Therefore, this question and its negative answer seem to constitute a QA contradiction.

Generally speaking, asking a question usually involves the expectation that an addressee is talking sincerely. To induce an addressee to speak sincerely is part of the *perlocutionary act* of questioning. Generally, an *illocutionary act* can succeed even if it does not result in a perlocutionary

⁴ In my “Question–Answer Contradiction,” I argued that the question constituting a QA contradiction is in a pragmatic contradiction. However, this was not precise, and such a question is actually *nonconforming*.

act. However, the questioner's intent also includes the perlocutionary act. Therefore, questioning necessarily involves the expectation that an addressee will answer sincerely. In this case the questioner assumes the following two conflicting attitudes: (i) the questioner doubts whether an addressee is talking sincerely, but (ii) the questioner intends to obtain a sincere response from the addressee. These two attitudes may not be contradictory, but they are nonconforming.

In this case the negative answer "I'm not talking to you sincerely" involves a pragmatic contradiction, similar to utterances such as "I am a liar" or "I am joking." However, the negative answer "I am not talking to you sincerely" involves not only a pragmatic contradiction but also a QA contradiction. The utterance is a response to a particular question. Responding to a question generally requires a sincere attitude to the question. (We can make a joke, but in order to succeed in joking, we must convey a metamessage that the message is a joke by means of voice, intonation, nonverbal behaviour, or the like. The joke would work only as part of a conversation that is sincere as a whole. Responding to being addressed usually requires a sincere response.) Therefore, the content of the negative answer and the responding relation of the answer to the question are contradictory. This type of QA contradiction is *mixed* because it involves a pragmatic contradiction in the answer.

(2) *Definition of mixed QA contradictions.* I label the QA contradiction exemplified above a *mixed* QA contradiction, which is one in which the answer involves a pragmatic contradiction. Other examples of this type include the following:

"Are you a liar?"—"Yes, I am a liar."

"Does anyone exist?"—"No, nobody exists."

"Do you exist?"—"No, I don't exist."

"Do I exist?"—"No, you don't exist."

"Are you claiming anything?"—"No, I am not claiming anything."

"Does one need grounds for making an assertion?"—"No, one need not have grounds to make an assertion."

"Should I obey your order?"—"No, you should not obey my order."

"Who is there?"—"Nobody is here."

"Which language can you use?"—"I cannot use any language."

All of the above examples of *pure* QA contradiction and *mixed* QA contradiction demonstrate that a QA contradiction involves fundamental aspects of communication or *relations of linguistic mutual responsiveness*. However, the three examples explained in detail above are especially important examples, because they demonstrate that the confirmation of connection, of understanding a language, and of sincerity are fundamental conditions for communication. In the following sections, I attempt to persuade the reader further on this point, first by offering a general analysis of QA relations with respect to anaphora and presuppositions of questions, and then by applying this framework to QA contradictions.

Analysis of QA Relations

Anaphora and QA Relations

(1) *Anaphora in QA relations.* Anaphora is the use of an expression (an anaphor or a dependent) to refer to another previous expression (an antecedent), both of which refer to the same object. For example, in the sentence “This organism is a mammal, therefore it is a vertebrate,” the word “it” refers to “this organism” in some sense; more precisely, it does not refer to the expression itself, but to the object to which “this organism” refers. The expression “it” is parasitic on the antecedent expression “this organism.” An anaphoric expression could be said to have *parasitic* reference or *indirect* reference. This type of reference may at times involve an anaphoric chain.⁵

Anaphora is inevitable in QA relations. In order to answer a question in the form of a complete sentence, an answer must repeat expressions that occurred in the question or refer to them anaphorically through the use of pronouns. In this sense, every answer contains some anaphoric expressions. For example: “What is this apple?”—“It is a Macintosh

⁵ On this point I have benefited greatly from personal communication with Robert Brandom. He emphasizes the importance of anaphora in *Making It Explicit: Reasoning, Representing, and Discursive Commitment* (Cambridge, MA: Harvard University Press, 1994).

apple.” Another example: “Is her friend a philosopher?”—“Yes, she is a philosopher.” The pronouns “it” and “she” in these answers are anaphoric dependents. In some cases, coordination is needed in order to establish the appropriate dependent in the answer, illustrated here as follows: “Which is your car?”—“My car is that red car.” The phrase “my car” in the answer refers to the same object as “your car” in the question; hence “my car” is anaphorically dependent on “your car.”⁶ But anaphoric expressions such as these are often omitted from the answer, for example: “Which is your car?”—“That red car.” Or “What is this apple?”—“A Macintosh apple.” The anaphoric expressions in these answers represent already-known information. Insofar as an utterance is understood to be an answer to a particular question, the omission is evident. Therefore, we omit such anaphoric expressions relatively often.

The next example is an anaphoric relation between different languages: “Can you speak English?”—“*Ja, ich kann English sprechen.*” In this case, “you” in the question and “*ich*” in the answer stand in an anaphoric relation. Incidentally, as Davidson pointed out, we can communicate without using a common language.⁷ Therefore an anaphoric relation can hold between different languages.

(2) *Deictic use of token-reflexive expressions presupposes an anaphoric mechanism.* We often use *indexicals* in QA relations, as in the above examples, with “I” and “you” playing particularly important roles. So let us think about the relation between the indexical “I” and anaphora. “I” is usually defined as follows: “I” refers to the speaker or writer of the relevant occurrence of the word “I.”⁸ The second “I” in this definition is

⁶ In general, a question asks the listener to refer to an object, and an answer refers to that object. To do that, a question must refer to the object using some expression, and the answerer provides another expression of the object. These two expressions are coreferential, but differ in the ways in which the object is referenced. Cf. my “Identity Sentences as Answers to Questions,” *Philosophia Osaka* 7 (2012/13), 79–94, http://ir.library.osaka-u.ac.jp/dspace/bitstream/11094/23292/1/po_07-079.pdf.

⁷ When X can understand German but cannot speak German, and Y can understand English but cannot speak German, X and Y can communicate with each other without using a common language. Donald Davidson pointed this out in “The Second Person,” in *Subjective, Intersubjective, Objective* (Oxford: Oxford University Press, 2001), 107–21.

⁸ David Kaplan, “Demonstratives,” in *Themes from Kaplan*, ed. Joseph Almog, John Perry, and Howard Wettstein (Oxford: Oxford University Press, 1989), 505.

an anaphoric use, and its antecedent is the first “I”—that is, the second “I” is referentially dependent on the first “I.” If the first “I” and the second “I” refer to different persons, this cannot be a correct definition of “I.” Thus, we need an anaphoric mechanism for defining the indexical “I.” Specifically, indexicals such as “I” are *token-reflexive* expressions, which are defined as expressions whose reference is determined in relation to the particular token of that expression; for example, “I” refers to the person who has uttered that token of “I.” Note that the definition of token-reflexive expressions involves an anaphoric relation, hence the *deictic* use of “I” presupposes an anaphoric mechanism, as does the deictic use of other indexicals such as “you,” “here” and “there,” “now” and “then,” and “this” and “that.”

(3) *Substitutability of the performative use of “I” in QA relations.* The performative utterance, as identified by J. L. Austin, is thought to involve a sentence form in which the subject is a first person (singular or plural) pronoun, the tense is the present, and the verb is a performative verb. This is by and large correct. But I would like to point out that, in such cases, the first person pronoun can be replaced by other expressions in QA relations via anaphora.

“I” has performative uses, as illustrated here: “I order udon” (Japanese noodles). With regard to such performative uses, it is generally supposed that we cannot substitute coreferential descriptions for “I” because this would change the performative use of the original utterance to a *constative* (descriptive) use as follows: “The first guest today orders udon.” But such a substitution maintains performativity in the following specific kind of case: if the following answerer is the referent of “the first guest today,” her answer is performative, as in this question–answer pair: “What are you ordering?”—“The first guest today orders udon.” She can also perform the same act of ordering by using her proper name, Tomoko: “What are you ordering?”—“Tomoko orders udon.” In these cases, two factors are responsible for the performative nature of these answers. One is that “the first guest today” and “Tomoko” are anaphorically dependent on “you” in the question. These anaphoric relations are based on QA relations, or QA relations require these anaphoric relations. The second factor is that the question “What are you ordering?” requires a performative utterance as its answer.

(4) *A borderline case of anaphora in QA.* As we saw above, every QA relation involves anaphora. If that is the case, what should we make of the following example? Question: “Does anybody know it?”—Answer: “Hi!” This answer seems to involve no anaphoric mechanism. This might be understood as an abbreviated answer. First I would like to consider the relation between abbreviation and anaphora.

Consider this question–answer pair: “Do you finish it today?”—“Yes, I do so.” The “do so” in the answer could be understood as an anaphor or as an abbreviated answer. If it is an abbreviated answer, a full answer will be “Yes, I finish it today.” This full answer is anaphoric, because “it” is an anaphoric dependent and “finish” is also anaphoric: in order to understand the meaning of “finish” in the answer, we need to refer to the meaning of “finish” in the question due to the ambiguity of the word. Thus an abbreviated answer is also anaphoric. If the answer “Hi!” is an abbreviated answer, its full answer would be “Hi, I know it.” The “know it” in this answer can be understood as an anaphor of “know it” in the question. This case of an abbreviation of anaphoric expression is a borderline case of anaphoric expression.

By the way, can the answer “Hi!” be a non-abbreviated answer? If “Hi!” is an answer, it must be an abbreviated one, because the answers to this yes-or-no question can be of only two types: “Yes, somebody knows it” or “No, nobody knows it.” Therefore if “Hi!” is an answer, it must be one of these. It is very plausible that “Hi” means “Hi, I know it.” But it is logically possible that “Hi!” means “Hi, nobody knows it.” Also in that case “Hi” is an abbreviated answer and anaphoric. If we cannot decide which answer “Hi” means, “Hi” cannot be an answer to the question for us. This is a borderline case between an answer and a mere response anaphoric relation.

Linguistic communication involves a kind of relation of mutual responsiveness. In order to respond to X who addresses Y, Y must specify X, understand that an utterance by X is addressed to Y, and respond to X’s address to Y. Hence, Y’s response to X is simultaneously referring indirectly to Y. The token reflexivity of “I” might come from the responsive relation of an answer to a question. In the above case, if the “Hi!” is an abbreviation of “Hi, I know it,” the response “Hi” would refer to the questioner and indirectly to the responder him or herself.

So I could suppose that when the indexical “I” arose from relations of mutual responsiveness, it might have been primarily a word for responding, not for referring. In this context we could interpret Anscombe’s claim that “I” is not referential.⁹ “I” and “you” would have arisen from the mutual responding relation, and this might suggest to us that the expression “I” was introduced for the purpose of responding.¹⁰

(5) *Application of (3) to QA contradictions.* By applying the above analysis (of the substitutability of the performative use of “I” in QA relations) to a QA contradiction, we can see that anaphoric relations are fundamental for a QA contradiction. Here is an example. Tom: “Can you hear me?”—Jerry: “No, I can’t hear you.” If this question is uttered in a context, then we can derive its content; the content of the above exchange uttered in the same context is the same as the following exchange. Tom: “Can Jerry hear Tom?”—Jerry: “No, Jerry can’t hear Tom.” The latter question–answer pair also represents a QA contradiction, because “Jerry” and “Tom” in the answer must be anaphoric dependents if Jerry’s utterance is to qualify as an answer, though Jerry cannot talk anaphoric-dependently because he cannot hear the question. We could say the same thing about other QA contradictions in which “I” and “you” are used. This shows that a QA contradiction does not necessarily depend on using indexicals but rather on an anaphoric mechanism.

(6) *Application of (4) to QA contradictions.* The QA relation presupposes an anaphoric mechanism, because an answer in its complete form must repeat some parts of the question. Therefore, anaphora seems to be a transcendental condition of the QA relation. The following was a borderline case between an answer and a mere response. Q: “Does anybody know it?”—A: “Hi!” In this case, the negative answer “no” or “nobody” does not constitute a QA contradiction, but the next example does constitute one. Q: “Is anybody here?”—A: “Hi!”; or alternatively, “No!”

⁹ G. E. M. Anscombe, “The First Person,” in *Metaphysics and the Philosophy of Mind, Collected Philosophical Papers Vol. II* (Oxford: Blackwell, 1981), 21–36.

¹⁰ This could explain the fact that there are numerous first-person pronouns and second-person pronouns in Japanese and that the use of them depends on the social relation between interlocutors.

In this example, “Hi!” and “No!” contain no explicit anaphora. But as we saw above, we could interpret these as abbreviated answers. The negative answer “No!” here is a QA contradiction because, on the one hand, “No!” must be a response in order to be an answer to the question, but on the other hand, it cannot be a response because the responder does not exist. In this case, every possible answer or response becomes the same as the positive answer “Hi!” and a significant distinction in this situation is a distinction between a response and no response. The same applies in the following case. Q: “Can you hear me?” A: “Hi!”; or alternatively, “No.” The short answer of “Hi” might implicitly indicate an anaphoric relation to the question. But the negative answer “No” would construct a QA contradiction.

Presuppositions of Questions

The notion of presupposition can be viewed from the perspective of the *semantic* presupposition, as argued by Strawson, or the *pragmatic* presupposition, as advanced by Stalnaker. But we cannot extend either of these directly to questions.¹¹ When it comes to defining the presuppositions of questions, Nuel Belnap offers the following: “A question Q presupposes a statement A, if the truth of A is a necessary condition for the existence of a true answer to Q.”¹² I would like to accept this as a definition of the *semantic* presuppositions associated with questions,

¹¹ P. F. Strawson, “On Referring,” *Mind* 59, no. 235 (1950): 320–44; Robert Stalnaker, “Presuppositions,” *The Journal of Philosophical Logic* 2 (1973): 447–57; Robert Stalnaker “Pragmatic Presuppositions,” in *Semantics and Philosophy*, ed. Milton K. Munitz and Peter Unger (New York: New York University Press, 1974), 197–214. Stalnaker defends the notion of “the pragmatic presupposition” of a speaker against the semantic presupposition of a sentence. The concept of cognitive presupposition that I introduce here is similar to Stalnaker’s “pragmatic presupposition”; however, I suppose that semantic presuppositions of a question and cognitive presuppositions of it are compatible, and Stalnaker emphasizes the knowledge shared between two speakers. In this chapter, I focus on the differences in cognitive presuppositions between questioners and answerers.

¹² Cf. Nuel D. Belnap Jr and Thomas B. Steel Jr, *The Logic of Questions and Answers* (New Haven, CT: Yale University Press, 1976), 108–25; Henry Leonard, “Interrogatives, Imperatives, Truth, Falsity, and Lies,” *Philosophy of Science* 26, no. 3 (1959): 172–86.

and to introduce here another concept related to presupposition, that of *cognitive* presupposition.

(1) *Semantic and cognitive presuppositions of questions.* Consider the following question Q1: “Is this bird a migrant bird?” Q1 presupposes propositions such as the following:

SP1. This bird is either a migrant bird or not.

SP2. The referent of “this bird” exists.

SP3. The referent of “this bird” is a bird.

Necessarily, if Q1 has a true answer, then these propositions must be true. I will refer to such propositions as *semantic* presuppositions of questions. If a semantic presupposition of a question does not hold, then the question cannot have a true answer and is an invalid question. I will refer to a question that has a true answer as a *valid* question, and an utterance that, if true, becomes a true answer, as a *possible* answer. Asking a question amounts to dividing all utterances into possible answers versus others.¹³ Every possible answer must meet the semantic presuppositions of the question.

Q1 also comes with presuppositions of a different kind, such as the following:

CP1. The questioner knows which object “this bird” refers to.

CP2. The questioner understands the meaning of “a migrant bird.”

CP3. The questioner does not know the answer to the question.

Necessarily, if one asks Q1 sincerely, these propositions hold. I will call such propositions *cognitive* presuppositions of the questioner. If the questioner of Q1 does not meet CP1 or CP2, he or she cannot ask Q1 sincerely. There exist other such cognitive presuppositions in addition to CP1 and CP2. CP3 is somewhat different from the other

¹³This definition does not cover practical questions that call for performative utterances without truth values in response. I cannot deal with semantic or cognitive presuppositions of practical questions in this chapter because of limitations of space.

cognitive presuppositions. If CP3 is not met, then the questioner knows the answer to the question and need not ask Q1. By the way, even if the questioner of Q1 fails to meet some or all of the cognitive presuppositions, Q1 can have a true answer. In this sense cognitive presuppositions are different from semantic presuppositions. This is a criterion for distinguishing between the two kinds of presupposition.

If P is a semantic presupposition of a question, then it is a cognitive presupposition that the questioner believes P. Therefore, we can obtain the following cognitive presuppositions respectively from the above three semantic presuppositions:

CP4. The questioner believes that this bird is either a migrant bird or not.

CP5. The questioner believes that the referent of “this bird” exists.

CP6. The questioner believes that the referent of “this bird” is a bird.

We should use “believes” here, not “knows,” because “knows” is too strong. Recall this famous example: the question “Is the present king of France bald?” has the semantic presupposition that there is a present king of France. The person who asks this question has the cognitive presupposition that he or she *believes* that there is a present king of France. He or she cannot *know* that there is a present king of France, because that is not the case.

Consider the question Q1 again: “Is this bird a migrant bird?” Q1 can be asked in different contexts or for different purposes. In a typical case, one might ask Q1 to find out about the bird referred to as “this bird.” But one might also ask Q1 to find out the meaning of “migrant bird.” In this case, the questioner would know a good deal about the bird referred to as “this bird.” In the second case, CP2 as identified above is not a cognitive presupposition of the questioner. In this case, CP7 is a cognitive presupposition:

CP7. The questioner believes that he knows a good deal about the bird referred to as “this bird.”

In the second case, if the questioner did not know the crucial properties of the bird referred to as “this bird,” he or she would not ask Q1. But the

questioner in the first case does not meet CP7. Thus, the cognitive presuppositions of the questioner vary depending on the context or purpose for asking the question.

Semantic presuppositions do not vary across contexts. In normal cases, the questioner of Q1 must believe its semantic presuppositions; otherwise he or she is asking Q1 without believing that it is a valid question. Therefore, belief in the semantic presuppositions of a question is always included among the cognitive presuppositions. This type of cognitive presupposition is common across all contexts in which Q1 is asked, because semantic presuppositions do not vary across contexts.

Note also that the order of questioning is changeable, as the following example illustrates. Q1: “Is this bird a migrant bird?”—Q2: “What kind of bird is a migrant bird?” If one asks Q1 to find out whether the bird in question is migratory, the following hold as cognitive presuppositions:

- CP1. The questioner knows which object “this bird” refers to.
- CP2. The questioner understands the meaning of “migrant bird.”
- CP3. The questioner does not know the answer to the question.

In this case, the questioner does not ask Q2 after asking Q1, because this would contradict CP2. If one asks Q1 so as to know the meaning of “migrant bird,” then the following hold as cognitive presuppositions:

- CP1. The questioner knows which object “this bird” refers to.
- CP7. The questioner believes that he or she knows a good deal about the bird referred to as “this bird.”
- CP3. The questioner does not know the answer to the question.

In this case, the questioner could ask Q2 after asking Q1 to make his or her understanding of “migrant bird” more clear.

(2) *Relation between presuppositions and the answer or answerer.* First, I would like to clarify the relation between a possible answer and the *semantic* presuppositions of a question. In order for a claim to be a possible answer to a question, it must logically imply all semantic presuppositions of the question. This is a necessary condition of possible answers, but it is not a sufficient condition because, for example, the

claim that “this bird is either a migrant bird or not” implies all semantic presuppositions of Q1 but is not a possible answer.

If a claim is incompatible with a single semantic presupposition of a question, it cannot be an answer to that question. For example, the claim “There is presently no present King of France” is incompatible with a semantic presupposition of the question “Is the present King of France bald?” thereby making the question invalid. Therefore, it cannot be an answer to the question. But the claim can be a *sincere useful response* to the question. To answer a question, one must be committed not only to his or her answer but also to all semantic presuppositions of the question.

Second, I would like to clarify the relation between a possible answer and the *cognitive* presuppositions of a question. In order for an addressee to answer a question, she must acknowledge all cognitive presuppositions of the questioner; otherwise she must be in doubt about the sincerity of the questioner, as indicated in the following exchange: “Is there a Santa Claus?”—“I think you know the answer.” This claim denies a cognitive presupposition of the question, thereby casting doubt on the sincerity of the questioner; it does not constitute an answer to the question. But this claim could be a *sincere useful response* to the question without denying its validity. This is an important difference between the denial of semantic presuppositions and the denial of cognitive presuppositions.

In some cases, a responding claim denies both kinds of presupposition, as in the following example: “Have you stopped beating your wife?”—“Stop joking, you know I’ve never beaten my wife.” But we can still acknowledge that this claim may be a sincere useful response to the question.

(3) *Application of (1) and (2) to QA contradictions.* To begin with, let us consider an example of a pure QA contradiction. Q: “Can you hear me?”—A: “No, I cannot hear you.” The semantic presuppositions of question Q include the following:

The addressee referred to as “you” exists.

The questioner referred to as “me” exists.

There is a possibility of the addressee being able to hear the question.

The cognitive presuppositions of the question include the following:

The questioner believes all semantic presuppositions.

The questioner does not know the answer to the question.

Does the propositional content of the negative answer imply all semantic presuppositions? This is a difficult question to answer. We might think that the negative answer implies all semantic presuppositions. Even if this were so, the answer A cannot be a possible answer to Q. An utterer of A cannot hear the question; therefore, he or she cannot acknowledge the cognitive presuppositions of the questioner and simultaneously cannot doubt the sincerity of the questioner.

Second, let us consider an example of a *mixed* QA contradiction: “Are you talking to me sincerely?”—“No, I’m not talking to you sincerely.” The semantic presuppositions of this question include the following:

The addressee referred to as “you” exists.

The questioner referred to as “me” exists.

The addressee is talking to the questioner.

The cognitive presuppositions of the question include the following:

The questioner believes all semantic presuppositions.

The questioner does not know the answer to the question.

The content of the negative answer is not incompatible with any of the semantic presuppositions, and the answerer might acknowledge the cognitive presuppositions and the sincerity of the questioner. In this case, it is difficult to explain a QA contradiction based on the relation between the answer and the semantic or cognitive presuppositions of the question. Why is this the case?

The QA contradiction seems to arise at a more fundamental level. As I explained in the first section, a QA contradiction is a contradiction between the propositional content of an answer and the responding relation of the answer to a *nonconforming* question. In the example of a *pure* QA contradiction, the propositional content of the answer

(namely, that the answerer cannot hear the questioner) contradicts a normal QA relation (namely, that the answerer can hear the question). In the example of a *mixed* QA contradiction, the proposition of the answer (namely, that the answerer is not speaking sincerely) contradicts a normal QA relation (namely, that the answerer is responding sincerely to the questioner).

These conditions reflecting normal QA relations are usually contained within the set of semantic or cognitive presuppositions. But these conditions, namely, that an answerer can hear a question and that an answerer is responding sincerely, are not included among the presuppositions of the above questions, which are nonconforming in this sense. Therefore, the above answers are compatible with the questions' semantic and cognitive presuppositions. The questions that cause QA contradictions are nonconforming, that is, they lack the normal semantic and cognitive presuppositions.

The QA contradiction reveals the conditions that underlie normal fundamental QA relations, for example that the questioner and the answerer can hear each other, and that the questioner and the answerer are speaking sincerely. Every example of QA contradiction highlights a normal fundamental QA relation.

Transcendental Arguments Based on QA Contradiction

According to Robert Stern, “as standardly conceived, transcendental arguments are taken to be distinctive in involving a certain sort of claim, namely that X is a necessary condition for the possibility of Y —where then, given that Y is the case, it logically follows that X must be the case too.”¹⁴ What is often used in

¹⁴ Robert Stern, “Transcendental Arguments,” in *The Stanford Encyclopedia of Philosophy* (Summer 2015 edition), ed. Edward N. Zalta, <http://plato.stanford.edu/archives/sum2015/entries/transcendental-arguments/>.

transcendental arguments to show the *necessity* of X as a condition for the possibility of Y would be a *logical*, *semantic*, or *pragmatic* contradiction. I would like to make a claim here about the necessary conditions for the possibility of the QA relation by using a QA contradiction. This would be an alternative kind of transcendental argument.

A QA contradiction is one that arises when we violate the fundamental conditions for QA relations. Therefore, the kind of answer that is necessary for avoiding a QA contradiction reveals something about necessary conditions for the possibility of QA relations. Probing QA contradiction allows us to identify the necessary or *transcendental* conditions for possible QA relations.

R. G. Collingwood claimed the following in *An Autobiography*: “In order to find out his meaning you must also know what the question was (a question in his own mind, and presumed by him to be in yours) to which the thing he has said or written was meant as an answer.”¹⁵ I acknowledge this claim by Collingwood.¹⁶ If it is correct, then the QA relation would be the core of our relations of linguistic mutual responsiveness. Thus, the transcendental conditions for possible QA relations would be simultaneously the transcendental conditions for the possibility of linguistic communication.

The basic relations of mutual responsiveness that define possible QA relations would be at the core of communication, that is, they represent *relations of linguistic mutual responsiveness*. Therefore, the conditions that define possible QA relations would be at the core of conditions that permit the possibility for relations of linguistic mutual responsiveness. What presuppositions accompany relations of linguistic mutual responsiveness? Relations of

¹⁵R. G. Collingwood, *An Autobiography* (Oxford: Clarendon Press, 1978), 31. He also says, “Every statement that anybody ever makes is made in answer to a question.” See his *An Essay on Metaphysics* (Oxford: Clarendon Press, 1940), 23.

¹⁶I tried to prove this in my paper “A Proof of Collingwood’s Thesis,” *Philosophia Osaka* 4 (2009): 69–83, http://ir.library.osaka-u.ac.jp/dspace/bitstream/11094/10747/1/po_04_069.pdf.

linguistic mutual responsiveness appear to presuppose at least the following conditions:

1. Interlocutors can hear each other's voices (*the mutual confirmation of connection*).
2. Interlocutors can understand each other's language (*the mutual understanding of language*).¹⁷
3. Interlocutors speak sincerely to each other (*the mutual confirmation of sincerity*).

The following QA contradictions make it explicit that these three are necessary conditions for QA relations.

“Can you hear me?”—“No, I cannot.”

“Can you understand my language?”—“No, I cannot.”

“Are you speaking sincerely?”—“No, I am not.”

These QA contradictions reveal that, if asked, addressees necessarily answer “Yes, I can hear you,” “Yes, I can understand your language,” and “Yes, I am speaking sincerely.” I will call such necessities *QA necessities*. These QA necessities indicate that whenever QA relations hold equally between two persons, conditions (1), (2), and (3) above necessarily hold. Therefore (1), (2), and (3) would be at least basic conditions that define QA relations.

Semantic Conditions for QA Relations

As explained above, anaphora is a semantic condition that makes a QA relation between a question and answer possible. The following QA contradiction makes this explicit. “Can you use anaphora?”—“No, I

¹⁷As Davidson pointed out in “The Second Person,” they need not share a common language.

can't use it." The "it" in the answer is anaphoric. The "use" in the answer is also anaphoric, because "use" is ambiguous and its meaning in the answer needs to be understood as being the same as the meaning of "use" in the question. The linguistic response is principally anaphoric. This QA relation is in a QA contradiction between the propositional content of the negative answer and the responding relation of the answer to the nonconforming question.

As noted above, the following case must be explained: "Does anybody exist?"—"Hi!" If we can interpret this answer as an abbreviated expression of "Yes, I exist," then we can say that an anaphoric expression is omitted here, which is indicated by the responding relation between the question and the answer. So anaphoric relations and QA relations presuppose each other.

The distinction between *types* and *tokens* is also a necessary condition for the possibility of linguistic communication. If an expression is repeated anaphorically in two separate instances, this results in two occurrences of the same thing as two tokens of the same type. Hence, to distinguish types from tokens, we need to be able to repeat an expression. Anaphoric repetition and the distinction between types and tokens presuppose each other. QA relations presuppose anaphoric relations. Anaphoric relations imply, at least implicitly, repetitions. Therefore, QA relations presuppose the distinction between types and tokens. We can demonstrate this with the following example. "Can you *repeat* any word?"—"No, I cannot *repeat* any word." This is a QA contradiction. This means that whenever a QA relation holds, we are distinguishing between types and tokens.

Following linguistic rules is also a transcendental condition for the possibility of a QA relation. This can be shown by the following QA contradiction. "Can you repeat any words in order?"—"No, I cannot." Another example: "Can you follow any linguistic rules?"—"No, I cannot." These necessary semantic conditions for QA relations are justified by certain QA contradictions, because violations of such semantic conditions constitute QA contradictions. These semantic transcendental conditions could be proven by other means, but they can be proved by QA contradictions.

Logical Conditions for QA Relations

The *law of identity* presupposes anaphora. The law of identity can be expressed by the sentence “A is A.” We can justify “A is A” by using a QA contradiction as follows: “Is A A?”—“No, A is not A.” This negative answer represents a QA contradiction, because the first “A” in the answer is an anaphoric dependent of the first “A” in the question; therefore, the identity relation holds between both of these occurrences of “A,” as expressed in “A is A.” The responding relation of the answer to the question depends on an anaphoric relation, which is expressed in “A is A.” Therefore, this responding relation contradicts the content of the negative answer, resulting in a QA contradiction. So, if asked “Is A A?” the addressee necessarily answers “Yes, A is A.” We can justify the law of identity not by the law of contradiction but by this *QA necessity*, that is, by appealing to a transcendental argument based on QA contradiction.¹⁸

Obligation to Speak With Grounds

One might argue as follows. The assertion “One can assert something without grounds” is a pragmatic contradiction, because the concept of the illocutionary act of assertion seems to require that to assert something is to assert it with grounds; thus, the propositional content of the assertion “One can assert something without grounds” and the illocutionary act of it are in contradiction. This argument presupposes that

¹⁸I would like to comment briefly on the difference between “Hesperus is Hesperus” and “Hesperus is Phosphorus.” We can explain the difference in terms of QA relations. “Hesperus is Phosphorus” can be the answer to a *what* question, but “Hesperus is Hesperus” cannot, because if one is asked “What is Hesperus?” one does not answer “Hesperus is Hesperus.” “Hesperus is Phosphorus” can be the answer to a *yes-no* question such as “Is Hesperus Phosphorus?” but “Hesperus is Hesperus” cannot, because the answer is evident and a question such as “Is Hesperus Hesperus?” cannot be asked sincerely. This suggests that “Hesperus is Phosphorus” has cognitive significance due to its status as an answer in a QA relation. Inferential semantics might attempt to explain the difference between these two sentences by the difference in their inferential roles. This difference can also be explained in terms of QA relations.

grounds are needed in order to assert something. Therefore, this line of argumentation cannot prove the obligation to speak with grounds.

However, we can appeal to QA contradiction here. Consider the following question and answer: “Should one have grounds when asserting something?”—“Yes, one should”; or alternatively, “No, one need not have grounds to assert something.” This negative answer can be identified as a QA contradiction based on the following. A question that requires an assertive utterance as its answer normally requires an answer with grounds, because an answer without grounds has no significance for the questioner. The above negative answer falls into QA contradiction between the propositional content of the answer and the conditions for a normal QA relation. Therefore, if we answer the question, we must reply with an affirmative answer to the question. The obligation to speak with grounds can be justified by appealing to a QA contradiction.

If it is widely acknowledged that to claim something is to claim something with grounds, then to claim something without grounds is akin to telling a lie. In the following section, I justify a general prohibition against lying.

Prohibition Against Lying

The question “May one tell a lie?” can be answered in one of two ways: “Yes, one may tell a lie” or “No, one may not tell a lie.” The answer “no” prohibits the telling of a lie. This utterance is not problematic. But the answer “yes” is problematic for the following reasons.

As Kant pointed out,¹⁹ if lying were allowed, it would become impossible because of the following. To tell a lie *P* is to claim *P* while believing that *P* is false. But *P* could be true even if the speaker believes it to be false. Hence, telling a lie *P* and the truth value of *P* are indifferent to each other. If lying were allowed, then one’s claim *P* might be a lie, that is, there would be no difference between claiming that *P* is true and claiming that *P* is not true. In such a situation, it would be impossible to claim something. Furthermore, it becomes impossible not only to claim

¹⁹ G 4:402f.

but also to promise, order, declare, and so on; it becomes generally impossible to say something. So the claim “One may tell a lie” will make it impossible to say “One may tell a lie.” According to Kant, moral law should be universally accepted; however, if “One may tell a lie” is universally accepted in a society, then people come not to believe that others speak truthfully, and it becomes impossible both to tell a lie and to tell the truth in that society. Therefore, when we think about such a social process, we cannot admit “One may tell a lie” as a universal law.

But if we focus on a QA contradiction that is involved in this case, we can justify the prohibition against lying in another way, without thinking about such social processes. The question “May one tell a lie?” demands a true answer. But the answer “Yes, one may tell a lie” renders the demand for a true answer invalid. Hence, the content of this answer and the responding relation of the answer to the question are in contradiction with each other. Therefore, to answer the question in a valid way, we must offer a negative answer to the question. The prohibition against lying can be justified using QA contradiction.

To consider this matter further, suppose that a speaker tells a lie by saying “P”; then if she were asked the reason for P, she would respond by offering grounds she believes to be false. So, telling a lie amounts to making a claim based on mendacious grounds. We can justify a prohibition against speaking without grounds as well as against speaking based on mendacious grounds.

Obligation of Mutual Recognition

As we saw in the above discussion of mixed QA contradictions, when we are asked “Are you talking sincerely?” we answer always “Yes, I am talking sincerely,” because negative answers such as “No, I am not talking sincerely” or “No” constitute not only pragmatic contradictions but also QA contradictions. Further, when we ask a person “Are you talking sincerely?” his or her answer is necessarily “Yes” by the above argument. Therefore, we should say “You are talking sincerely.” From both sides, we should say “We are talking sincerely to each other.” This would constitute a relation of mutual recognition. This suggests to us

that mutual recognition would be a necessary condition of a linguistic mutual responding relation. However, mutual recognition seems to involve more than this.

Donald Davidson explained his “principle of charity” as follows: “charity is forced on us; whether we like it or not, if we want to understand others, we must count them right in most matters.”²⁰ We are supposed to regard the other as a reliable rational being in order to understand an utterance of the other. What would it mean to be reliably rational in this context? It would mean meeting necessary conditions for the possibility of linguistic mutual responsiveness. Consider: “Do you meet the necessary conditions for the possibility of linguistic mutual responsiveness?”—“No, I don’t.” This negative answer falls into a QA contradiction, because the content of the answer and the responding relation of the answer to the question are contradictory. Therefore, it is necessary for interlocutors to regard each other as meeting the necessary conditions of a linguistic mutual responding relation. This is a kind of necessary condition for communication. This would justify the obligation of mutual recognition.²¹

Conclusion

The aim of this chapter was to propose a method for answering the question: What are the fundamental conditions for relations of linguistic mutual responsiveness? The method I have proposed is a transcendental argument based on QA contradiction. First, I defined a QA contradiction as a contradiction between the propositional content of an answer and a responding relation between the answer and the nonconforming question, while distinguishing between pure and

²⁰ Donald Davidson, “On the Very Idea of a Conceptual Scheme,” in *Inquiries into Truth and Interpretation* (Oxford: Oxford University Press, 1984), 197.

²¹ The “mutual recognition” discussed here relates to the concept “*Aufforderung*” in Fichte and the concept “*Anerkennung*” in Fichte and Hegel. I interpreted them using the concept of “double bind” in my Japanese book *Doitsu Kannenron no Jissen Tetsugaku Kenkyu* [Studies in Practical Philosophy in German Idealism] (Kyoto: Kōbundō, 2001).

mixed QA contradictions. Second, I explained that every QA relation presupposes an anaphoric relation, because a responding relation between question and answer needs an anaphora. I also analyzed the semantic presuppositions of questions and the cognitive presuppositions of the questioner, as well as their relation to possible answers and the answerer. However, I concluded that the QA contradiction relates to a more fundamental level of QA relations. Next I provided a partial list of the conditions necessary for a QA relation by presenting transcendental arguments based on QA contradictions. I argued that the mutual confirmation of connection, the mutual understanding of the language of the interlocutors, and the mutual confirmation of sincerity are all basic mutual responding conditions for QA relations. I further argued that anaphora, the distinction between types and tokens, and following linguistic rules are fundamental semantic conditions for QA relations, and also that the law of identity is a logical condition for them. Moreover, I argued that the obligation to speak with grounds and the prohibition against lying are deontic necessary conditions for QA relations. Finally, I argued for the necessity of mutual recognition as a transcendental condition of a QA relation and a linguistic mutual responding relation.

I suppose that the transcendental conditions involved here are more fundamental than the semantic and cognitive presuppositions of questions: the former are universal conditions, but the latter are empirical and vary case by case. With regard to these transcendental conditions, at the outset I hoped to justify our knowledge and ethical principles based on QA contradictions, much like Apel, who tries to justify them based on pragmatic contradictions. But now I think that such foundationalism is impossible. I agree with Rorty, who does not completely reject transcendental arguments but admits them only as parasitic arguments that always presuppose some theory in order to prove the necessity of a conclusion.²² So I now think that transcendental arguments supply only a weak form of justification.²³

²² Cf. Richard Rorty, "Verificationism and Transcendental Arguments," *Noûs* 5, no. 1 (1971): 3–14.

²³ I appreciate the many useful comments Halla Kim offered on an earlier draft.

Yukio Irie is Professor at the Department of Philosophy, the Graduate School of Letters, Osaka University. Recent publications include an anthology (with K. Nagasawa), *The Whole Picture of Fichte's Wissenschaftslehre* [in Japanese] (Kōyōshobō, 2014). Among his papers are “Fichte and Semantic Holism,” in *Kant, Fichte and the Legacy of Transcendental Idealism* (Lexington Books, 2014), and “Question and Inference,” in *Begegnungen in Vergangenheit und Gegenwart* (Lit Verlag, 2015). Among his research interests are philosophy of language, German idealism, and the history of Japanese thought.

12

Consequences of the Transcendental-Pragmatic Consensus Theory of Truth

Michihito Yoshime

It can be said that the current trend in philosophy, especially with regard to theories influenced by analytic philosophy, hermeneutics, or postmodern thinking, is toward empiricism, realism, and fallibilism, and therefore toward anti- (or “post-”) transcendentalism. In this context, it is not surprising that the transcendental pragmatics proposed by the German philosopher Karl-Otto Apel¹ has not garnered broad acceptance. In fact, this linguistically and philosophically transformed transcendental philosophy (in the Kantian sense) has been

¹Primarily in Karl-Otto Apel, *Towards a Transformation of Philosophy*, trans. Adey Glyn and David Frisby (London: Routledge & Kegan Paul, 1980). This is an abridged translation of the original German edition from 1973. For German sources, I have consulted and cited published English versions whenever possible, although I have slightly changed some expressions. I take full responsibility for all other translations of German writings (which are cited by their German titles below).

M. Yoshime (✉)

Department of Philosophy, Sophia University, Tokyo, Japan
e-mail: michihito.yoshime@gmail.com

embroiled in ongoing disputes owing to its strong assertions, including its notion of the *ultimate grounding* of philosophical knowledge and norms in discourse ethics. Such grounding should be achieved through the *Unhintergebarkeit* (uncircumventability) of the conditions of our argumentative discourses. The point of the ultimate grounding consists in the requirement that we admit those propositions that we cannot deny without a performative self-contradiction. Apel provides the following definition: “If I cannot challenge something without actual self-contradiction and cannot deductively ground it without formal-logical *petitio principii*, then that thing belongs precisely to those transcendental-pragmatic presuppositions of argumentation which one must always have accepted, if the language game of argumentation is to be expected to retain its significance.”²

Apel presents the idea of the uncircumventability of argumentative discourses and their norms in his earlier works; here, he provides the exact meaning of this strange term.³ But many philosophers regard this argument as impossible to understand and consider the whole program to have failed.

It is arguable, however, that this contemporary transcendental approach is not only possible but also has many valid points.⁴ In this chapter, I will focus on another of its controversial assertions, the *consensus theory of truth*, which has also been the subject of numerous criticisms. I will address some pertinent criticisms, namely one by Albrecht Wellmer and another by Hilary Putnam. Both Wellmer and Putnam regard transcendental pragmatics as a kind of anti-realism, which is, to them, not acceptable, although my own opinion is that

² Karl-Otto Apel, “The Problem of Philosophical Fundamental-Grounding in Light of a Transcendental Pragmatic of Language,” trans. Richard Pavlovic, *Man and World* 8, no. 3 (1975): 264.

³ Cf. also Wolfgang Kuhlmann, *Reflexive Letztbegründung* (Freiburg and München: Karl Alber, 1985), 82–3.

⁴ I have pointed out the significance of transcendental pragmatics from the perspective of the rediscovery and rehabilitation of the Fichtean transcendental approach. See Michihito Yoshime, “What is the Unlimited Communication Community? Transcendental Pragmatics as Contemporary Fichteanism,” in *Fichte and Transcendental Philosophy*, ed. Tom Rockmore and Daniel Breazeale (Basingstoke, UK: Palgrave Macmillan, 2014), 273–92.

there is limited validity to their criticisms. However, rather than directly responding to them, I will leverage their criticisms for the purpose of clarifying the transcendental-pragmatic notion of truth and illustrating its relevance to ultimate grounding, with the end result that these criticisms will become less compelling.

First, I explain how Apel adopts and integrates two variations of the consensus theory of truth as articulated by Charles S. Peirce and Jürgen Habermas. Second, I briefly reconstruct Wellmer's criticism of these accounts, and point out a metaphysical-realistic presupposition that is implied by the criticism. Third, I investigate Putnam's understanding of Apel's view of truth, then link consensus theory with the confrontation between realism and anti-realism and identify Kant's position in this debate. Fourth, I contrast the metaphysical-realistic concept of truth with the transcendental-pragmatic one, and I point out a few advantages of the latter. Finally, I examine and clarify the meaning of "ultimate" grounding in light of a transcendental anti-realistic understanding of truth and identify several positive consequences that follow from that understanding.

What is the Consensus Theory of Truth?

The consensus theory of truth, as developed by transcendental pragmatics, regards truth as the ultimate consensus of the ideal communication community, building upon Peircean consensus theory and accepting Habermas's variation. In developing his theory, Apel was most strongly influenced by Peirce. One can readily discern the influence of Peirce's semiotic interpretation of Kant on his formulation of transcendental pragmatics. Peirce's consensus theory of truth, which is combined with the Kantian interpretation, is also one of the elements that determined the fundamental direction of Apel's philosophy.

Apel approaches the problem of an adequate explication of the meaning of truth with the notion that such an explication should be "criteriologically relevant."⁵ That is, it should be capable of specifying

⁵ Karl-Otto Apel, "C. S. Peirce and Post-Tarskian Truth," in *Selected Essays*, vol. 1, trans. and ed. Eduardo Mendieta (Atlantic Highlands, NJ: Humanities Press 1994), 178.

a criterion for truth, rather than of dealing with the meaning of its concept. Based on this idea, Apel regards Tarski's logical-semantic explication of the meaning of truth as criteriologically irrelevant. Tarski appears to offer a fine criterion for truth, namely correspondence to the facts, by introducing the convention T ("p" is true if, and only if, p). However, this definition is not limited to the correspondence theory of truth, because it is restricted within the frame of formal language and is, as Tarski himself emphasizes, neutral with respect to various ontological or epistemological theories (and therefore also with respect to various truth theories).⁶ On the other hand, there have been philosophical attempts to provide a criterion for truth, and these can be classified into several types. Among these, Apel designates the correspondence theory, the coherence theory, and the evidence theory as relevant; however, each of these is defective and, in its current state, not applicable to contemporary debate regarding the meaning and criterion for truth.

The classic correspondence theory, which was broadly espoused from Aristotle to the medieval metaphysicians, has the problem that we cannot confirm the actual correspondence between thought and reality. That is, if we want to avoid the "God's eye view," correspondence must be confirmed through thought and the appearance of reality, which also reflects mere thought. This leads to a regress ad infinitum. If one tries to avoid this, as do more modern versions of this theory, by simply describing a correspondence between propositions and facts that are conceived of as existing if and only if the corresponding propositions are true, then the result is a circular argument between correspondence and truth.⁷

⁶ Cf. Apel, "Peirce and Post-Tarskian Truth," 176–7. This becomes more obvious if the more recent "deflationary theory" of truth is considered. Paul Horwich accepts the equivalence schema (E): It is true *that* *p* if and only if *p*. However, he asserts that the truth predicate "exists solely for the sake of a certain logical need" to "adopt some attitude towards a proposition—for example, believing it, assuming it for the sake of argument, or desiring that it be the case," and that "we can say what is in the basic theory of truth—an infinity of biconditionals of the form of '*p*' is true if *p*'—but we can't formulate it explicitly because there are too many axioms [referring to correspondence, coherence, and so on]." See Paul Horwich, *Truth*, 2nd edn (Oxford: Clarendon Press, 2005), 6, 2, 11.

⁷ Cf. Apel, "Peirce and Post-Tarskian Truth," 184.

A simple way to avoid this problem is to consider that propositions (note that thoughts are also propositions) can only be compared with, and verified by, propositions; this leads to the coherence theory of truth, which is also unsatisfactory. Here, Apel focuses particularly on Lorenz B. Puntel's coherence theory, which was influenced by Hegel. This version regards coherence as "the ultimate standard or regulative principle of a possible integration of truth criteria,"⁸ thus providing a philosophical explication of the meaning of truth and, at the same time, offering a coherent version of all possible criteria as its ultimate criterion. However, according to Apel, such a "pure coherence theory"⁹ overestimates the importance of the notion of coherence. This is because, if we distinguish mere thought about possible worlds from our knowledge of the real world, we need to rely more on other truth criteria, including correspondence and evidence. The view that coherence is the only criterion for truth in every philosophical context is going too far.

Another option is what Apel calls the evidence theory of truth, the position represented by Husserl, who succeeded Descartes in this regard; this account replaces correspondence with evidence of fulfillment of our "meaning-intentions."¹⁰ Apel holds this account (which resembles Peirce's ideas, to a degree) in high esteem, although he also expresses reservations about its methodological-solipsistic method, which stems from its nature as a philosophy of consciousness. He repeatedly notes the insufficiencies of the Cartesian/Husserlian methodological solipsism. For example:

The certainty of the "*cogito, sum*" cannot, as E. Husserl wishes in *Cartesian Meditations*, be understood as being no longer formulable in the "communicative plural." This is because, in such an epoché of "methodological solipsism" in which the existence of other subjects would be bracketed along with the real world, the evidence of Cartesian insight [of *cogito*] in principle could not be formulated in the sense of an intersubjectively valid philosophical judgment.¹¹

⁸ Apel, "Peirce and Post-Tarskian Truth," 185.

⁹ Apel, "Peirce and Post-Tarskian Truth," 185.

¹⁰ Apel, "Peirce and Post-Tarskian Truth," 186.

¹¹ Apel, "Problem of Philosophical Fundamental-Grounding," 266.

However, it must be noted that Apel does not consider evidence to be totally irrelevant to truth. He merely emphasizes that this position is satisfactory only if it is presupposed that the linguistic interpretation of a given phenomenon must be shared (in principle, transcending all the differences that may exist among various life forms) among all who participate in the redemption of the truth claim (in Habermas's sense); that is, it relies on the presupposition of an unlimited communication community as an interpretation community, at least in principle.¹²

From the above, it is obvious that adequate criteria need to be integrated into a more comprehensive theory of truth.¹³ Here, Apel introduces the Peircean consensus theory of truth, a truth theory that can integrate all truth criteria addressed in the aforementioned three theories. Peirce's views regarding truth are summarized in the following passage, which Apel has frequently quoted: "The opinion, which is fated to be ultimately agreed to by all who investigate, is what we mean by the truth, and the object represented in this opinion is the real."¹⁴ However, Peirce's writings are fragmentary and, although suggestive of the idea of a consensus theory of truth, do not fully explicate the content of such a theory. Hence, Apel makes reference to the universal-pragmatic doctrine of Habermas, namely, the principles of the propositional-performative double structure of human speech,¹⁵ the four validity claims (that is,

¹² Cf. Karl-Otto Apel, "Transcendental Semiotics and Truth: The Relevance of a Peircean Consensus Theory of Truth in the Present Debate about Truth Theories," in *From a Transcendental-semiotic Point of View*, ed. Marianna Papastephanou (Manchester: Manchester University Press, 1998), 67.

¹³ Cf. Apel, "Transcendental Semiotics and Truth," 64.

¹⁴ Peirce, *Collected Papers*, 5:407. Citations from Peirce refer by volume and paragraph to: *Collected Papers of Charles Sanders Peirce*, vols 1–6, ed. Charles Hartshorne and Paul Weiss; vols 7–8, ed. Arthur W. Burks (Bristol: Thoemmes Press, 1998).

¹⁵ Under the influence of J. L. Austin, Habermas argues that our speech has two dimensions, namely the propositional component of the sentence *and* the performance of a speech act (illocutionary act in Austin's terminology), which can be expressed in the form of the performative sentence (such as "I assert that . . ."). When I make the utterance "snow is white," I am not only expressing the content of the sentence, but I am also performing a constative speech act. Although Austin made an additional distinction with respect to the performative side, namely between illocutionary and perlocutionary acts, Habermas does not consider this to be essential. Cf. Jürgen Habermas, *The Theory of Communicative Action*, vol. 1, trans. Thomas McCarthy (Boston: Beacon Press, 1984), 292.

of intelligibility, truth, rightness, and truthfulness (sincerity)), and the anticipation of the ideal speech situation.

According to Habermas,¹⁶ the truth of propositions is closely related to the truth claims of constative (assertive) statements: “the meaning of truth, therefore, can be explained only with reference to the pragmatics of a specific class of speech acts. What we mean by the truth or falsity of propositions can be shown only by examining the performance of constative speech acts.”¹⁷ Moreover, “the meaning of truth implicit in the pragmatics of assertions can be explicated if we specify what is meant by the ‘discursive redemption [*Einlösung*]’ of claims to validity. This is the task of the consensus theory of truth.”¹⁸ By means of the “discursive redemption” of validity claims, Habermas conceives of a possible rational consensus within the ideal speech situation.¹⁹

Apel refers to the community of participants involved in this ideal discourse as the *unlimited ideal communication community*. For Peirce, an ideal community is conceived with the progress of scientific knowledge in mind and is deemed to include “all who investigate,” but Apel enlarges this notion into a more general idea, encompassing even the domain of morality. The ideal community also serves as an interpretation community for the integration of all possible truth criteria, accomplished by means of argumentative discourse.²⁰ The consensus within the community is the “ultimate criterion or regulative principle” for such integration.²¹

¹⁶In this chapter I do not consider Habermas’s later position, which Apel criticizes as “the dissolution of Habermas’s language pragmatic discourse theory.” See Karl-Otto Apel, “Pragmatism as Sense-Critical Realism Based on a Regulative Idea of Truth: In Defense of a Peircean Theory of Reality and Truth,” *Transactions of the Charles S. Peirce Society* 37, no. 4 (2001): 446.

¹⁷Jürgen Habermas, *On the Pragmatics of Social Interaction*, trans. Barbara Fultner (Cambridge: Polity, 2001), 86.

¹⁸Habermas, *On the Pragmatics of Social Interaction*, 89.

¹⁹In response to a criticism, Habermas also states that the concept of the validity of a sentence cannot be explicated independently of the concept of redeeming the validity claim; Habermas, *Theory of Communicative Action*, 316.

²⁰Apel, “Peirce and Post-Tarskian Truth,” 197.

²¹Apel, “Peirce and Post-Tarskian Truth,” 194.

How Should Wellmer's Criticism be Read?

Albrecht Wellmer focuses on the discourse ethics that are common to Habermas's and Apel's positions and criticizes the consensus theory of truth that is presupposed therein. Wellmer begins by distinguishing between a consensus theory that provides a criterion for truth and one that only explains the meaning of truth. His criticism of the former is that a factual consensus, even a rational one (i.e., one that arises under ideal conditions), cannot be a criterion for truth.²² Wellmer is targeting Habermas's consensus theory here, and the latter, indeed, stated the following in his early work: "only a *grounded* consensus is a consensus we can reach in discourses. This one alone holds a truth criterion, but the meaning of truth is not the circumstance that a consensus will in fact be reached, but (the presupposition) that at any time and place, if only we enter into discourse, a consensus can be arrived at under conditions that warrant its being a grounded consensus."²³ In this argument, he apparently considers the consensus theory of truth to provide not only the explanation of the meaning of truth but also a criterion for it. That is, truth refers to a presupposed consensus under ideal conditions, and such grounding is the truth criterion for a consensus.

However, Habermas seems later to have changed his position and is, eventually, interested in explaining the meaning of truth rather than in defining the criterion for truth. He later offers the following statement regarding the relationship between the criteria for truth and the meaning of truth: "this [consensus] theory of truth provides only an explanation of meaning, it does not provide a criterion; in the end, however, it undermines the clear distinction between

²² Albrecht Wellmer, *The Persistence of Modernity*, trans. David Midgley (Cambridge: MIT Press, 1991), 160–2. Wellmer raises the example of the consensus among leading nineteenth-century physicists regarding the truth of Newton's theories. That consensus was surely rational, but the content of the consensus was not true.

²³ Jürgen Habermas, "Wahrheitstheorien," in *Vorstudien und Ergänzungen zur Theorie des kommunikativen Handelns* (Frankfurt: Suhrkamp, 1995), 160; translated by Apel in "Peirce and Post-Tarskian Truth," 191.

meaning and criterion.”²⁴ Hence, as Wellmer is aware, his criticism is not entirely valid. On the other hand, Habermas acknowledges that the fulfillment or nonfulfillment of conditions of validity (in this case, of truth) can be ascertained by means of the argumentative redemption of the validity claim under conditions that approximate the ideal speech situation.²⁵ Therefore, if Habermas regards the ideal speech situation as being capable of being approximated in reality, then argumentative redemption is also capable of such approximation; thus, Wellmer’s criticisms may apply to this ascertaining function.²⁶ This point could be expressed as follows: a factual discursive redemption of a validity claim, even a rational one (i.e., one that arises under ideal conditions), cannot be a criterion for truth.

The version of consensus theory that is limited to providing an explanation of the meaning of truth—Wellmer proposed that Apel’s consensus theory is a variant of this version—is criticized as follows. Not only is the idea of ultimate consensus in the ideal communication community empty (i.e., of no use for our understanding of what a grounded consensus is), but Apel also fails to notice a “dialectical illusion” (*Schein*)²⁷ in trying to identify *the absolute* as separated from historical context with the highest point within the historical world and bring it back into the continuum of history. In this sense, Apel’s view can be characterized as a form of absolutism.²⁸ In contrast, Wellmer suggests that discourse ethics require a more fallibilistic position, representing a weaker and more multidimensional means of grounding morality.²⁹

What can be noted about this criticism from Apel’s side? On the one hand, Wellmer seems to regard consensus even under ideal conditions as fallible after all, which means that it cannot serve as a truth criterion,

²⁴ Jürgen Habermas, *Autonomy and Solidarity*, rev. edn, ed. Peter Dews (London: Verso, 1992), 160.

²⁵ Habermas, *Autonomy and Solidarity*, 159–60.

²⁶ Wellmer, *Persistence of Modernity*, 165–7.

²⁷ Wellmer, *Persistence of Modernity*, 170.

²⁸ Wellmer, *Persistence of Modernity*, 180.

²⁹ Wellmer, *Persistence of Modernity*, 116.

and, on the other hand, he seems to hold that, if absolute truth is conceivable, then it must be separated from the historical world. These two points may appear to be independent of each other, but they have the same origin. That is, according to this point of view, any historical consensus cannot reach the truth *precisely because* the truth is cut off from our historical world. Considering *every* consensus as fallible amounts to nothing but worrying about the possibility that consensus fails to hit the mark, namely that it fails to achieve the status of a fact or reality that is completely independent of us, not only historically but also metaphysically. We can characterize this position as fallibilism based on metaphysical realism.³⁰ The following passage by Wellmer also suggests this position: “the fact that we hold something to be true with good reason cannot represent for us an additional reason for the truth of the thing we hold to be true.”³¹ Thus, if a proponent of this position wants to give a philosophical (criteriologically relevant) account of truth, he or she has to postulate the God’s-eye view as involving judgment about the correspondence between our consensus and things in themselves as *the absolute*, a position that is not persuasive within contemporary philosophical debates. Additionally, this consequence is precisely what Wellmer himself wants to reject.

It must, however, be added that Wellmer attempts to solve the problem by simply abandoning the notion of *the absolute*. That is, he defines “philosophical truth” in terms of each temporary hermeneutic arrival point that must constantly be renewed through the productive interpretation of texts,³² and suggests that this puts the matter right.³³ Hence, it may not be fair to ascribe to him a kind of metaphysical realism. Nevertheless, it is possible to perceive such a presupposition

³⁰ Among other scholars who take a similar position, we can name, for example, Karl R. Popper. As is well known, Popper places fallibilism at the center of his philosophy of science, but he nonetheless calls himself a metaphysical realist; Karl R. Popper, *Realism and the Aim of Science*, ed. W. W. Bartley III (London: Routledge, 1983), 80.

³¹ Wellmer, *Persistence of Modernity*, 162.

³² Wellmer, *Persistence of Modernity*, 179.

³³ Although I will not address this here, Apel offers a detailed criticism of Wellmer’s position. See Apel, “Pragmatism as Sense-Critical Realism,” 449–55.

behind his criticism of the consensus theory, insofar as a criteriologically relevant account of truth is at issue.³⁴ However, one might note that the same difficulty arises with regard to the transcendental-pragmatic idea of the ultimate consensus, whereas Apel explicitly claims that a satisfactory philosophical account of truth—of course, he considers his own account as such—should be criteriologically relevant. Therefore, to confront Wellmer's criticism, it seems important to clarify the metaphysical (or more precisely, the transcendental) implication of the idea of consensus and how it differs from that of things in themselves.

How Is the Consensus Theory Both Anti-realistic and Realistic?

Here I would like to introduce a secondary topic in proceeding with the argument. Hilary Putnam states that the consensus theory held by Peirce and Apel is a variant of anti-realism,³⁵ because "on such an account, it is metaphysically impossible for there to be any truths that are not verifiable by human beings."³⁶ I regard this suggestion as important for the argument I will advance in the present section, and I will return to Putnam's criticism of Apel later. Of present concern is the fact that we may be able to review Wellmer's criticism of the consensus theory in

³⁴ Of course, Wellmer himself does not claim such an account. By reducing the range of philosophical explication of truth to the temporary hermeneutic transition, he distances himself from the criteriologically relevant role of philosophy and is satisfied with this stance.

³⁵ Michael Dummett defines "realism" in a broad sense as follows: "statements of the disputed class [namely, about the physical world, mental events, mathematical objects, the past and the future, etc.] possess an objective truth value, independently of our means of knowing it: they are true or false in virtue of a reality existing independently of us." On the contrary, the belief that "a statement of the disputed class, if true at all, can be true only in virtue of something of which we could know and which we should count as evidence for its truth" is defined as "anti-realism." Putnam has this notion of anti-realism in mind. See Michael Dummett, *Truth and Other Enigmas* (Cambridge, MA: Harvard University Press 1978), 146.

³⁶ Hilary Putnam, *The Collapse of the Fact/Value Dichotomy and Other Essays* (Cambridge, MA: Harvard University Press, 2002), 123.

terms of the opposition between realism and anti-realism, addressing it on a metaphysical—or, more precisely, a transcendental—level.

Taking Putnam's suggestion into account, we may place the transcendental-pragmatic consensus theory of truth in opposition to the metaphysical-realistic assumption brought out by Wellmer. That is, we can regard the former as a kind of metaphysical (transcendental) anti-realism because transcendental pragmatics, following the Peircean idea, denies the presupposition of things in themselves as real (in the absolute sense) and explicates the meaning of "truth" as that which is ultimately agreed upon by the unlimited ideal communication community.

However, the matter is not that simple. Apel also applies the label "sense-critical realism"³⁷ to Peirce's position, which he views as the predecessor to his own. Thus, to develop our current account, we have to clarify what is meant by this label and explain its relation to transcendental anti-realism. What is sense-critical realism?

The term "sense-critical" (*sinnkritisch*) means that the particular argument, theory, or position is oriented to the criticism of the meaning (-lessness) of the opponent.³⁸ Apel understands the following assertion by Peirce to be sense-critical: "any truth more perfect than this destined conclusion [namely, which is fated to be ultimately agreed to by all who investigate] . . . any reality more absolute than what is thought in it, is a fiction of metaphysics."³⁹ According to this point of view, we cannot meaningfully contrast the ultimate consensus with a transcendent reality. In short, "sense-critical" as it is used here implies a criticism of the meaning (-lessness) of the metaphysical-realistic concept of truth.⁴⁰ The

³⁷ Apel, "Pragmatism as Sense-Critical Realism," 445.

³⁸ He uses the word "sense" (*Sinn*) in the way he uses "meaning" here, having nothing to do with Frege's distinction between *Sinn* and *Bedeutung*. In another work, Apel applies the term "meaning-critical" to Peirce's argument. See Karl-Otto Apel, "Transcendental Semiotics and Hypothetical Metaphysics of Evolution: A Peircean or Quasi-Peircean Answer to a Recurrent Problem of Post-Kantian Philosophy," in *Peirce and Contemporary Thought*, ed. Kenneth Laine Ketner (New York: Fordham University Press, 1995), 390.

³⁹ Peirce, *Collected Papers*, 8:12; supplement by Apel in "Peirce and Post-Tarskian Truth," 190.

⁴⁰ Apel also refers to his formulation of the philosophical ultimate grounding as sense-critical; Apel, "Problem of Philosophical Fundamental-Grounding," 264.

metaphysical reality is independent of human thoughts and statements. Thus, knowing about its estimated correspondence to our thoughts and statements must involve seeing it from a “God’s eye view,”⁴¹ as it is completely impossible for us to grasp, and it is “a fiction of metaphysics” to consider this to be possible. The transcendental-pragmatic account is oriented to the meaninglessness of the metaphysical-realistic concept of truth. Putnam’s understanding of this position as “anti-realism” is plausible, at least in this context.⁴²

How, then, is sense-critical realism realistic? Peirce refers to the realistic hypothesis that must be presupposed by the scientific method, all while presenting sense-critical arguments. This hypothesis asserts that “there are real things, whose characters are entirely independent of our opinion about them; . . . we can ascertain by reasoning how things really are and truly are.”⁴³ Scientific investigation is impossible unless scientists presuppose such a reality or truth and, furthermore, have a “great hope”⁴⁴ that they can approach it through truth criteria such as the coherence of theories and cognitive evidence that enlarges our empirical knowledge. Peirce thus admits truth and reality that are independent of our factual and empirical opinions. However, here it must be emphasized that reality and the truth cannot be considered as independent of us on a metaphysical level. Thus the Apel/Peirce theory of truth can be seen as realistic in a sense, while remaining anti-realistic as a whole.

However, if we regard this as a kind of realism, is it not simply a contradiction to assert that a position can simultaneously represent realism and anti-realism? Not necessarily. We have a good example of

⁴¹ Apel, “Pragmatism as Sense-Critical Realism,” 444.

⁴² As is well known, Putnam himself changes his position from metaphysical realism to internal realism and even to natural realism. I cannot here examine the detailed contents of each of these positions, but, in general, he makes the first shift to avoid the God’s-eye view (cf. Apel, “Pragmatism as Sense-Critical Realism,” 466–9), approaching the Kantian idea that the world is somehow dependent on the human mind. However, he comes to regard the “interface” between them as problematic and employs more naïve realism, which again admits truth independent of its rational acceptability to us. Hilary Putnam, *The Threefold Cord: Mind, Body, and World* (New York: Columbia University Press, 1999), part 1.

⁴³ Peirce, *Collected Papers*, 5:384.

⁴⁴ Peirce, *Collected Papers*, 5:407.

such arguments: the arguments that distinguish between empirical and metaphysical (or transcendental) levels and take an integral, reversible stance concerning truth and the real. For instance, Kant states in his well-known Fourth Paralogism in the first edition of the *Critique of Pure Reason* that transcendental idealism may be empirical realism and therefore represent a kind of dualism.⁴⁵

Transcendental idealism is the doctrine that “appearances are to be regarded as being, one and all, representations only, not things in themselves,” though it can empirically “admit the existence of matter.”⁴⁶ Therefore, according to this reversible position, we can admit the existence of external objects on the empirical level, while regarding the whole outer world as a mere phenomenon on the transcendental level. This shows us that insofar as we distinguish between transcendental and empirical viewpoints, and can consistently explain the relationship between these two dimensions, an anti-realism can be simultaneously viewed as a realism. Our temporary conclusion would therefore be as follows. A transcendental-pragmatic consensus theory of truth that succeeds the Peircean account is simultaneously a transcendental anti-realism and an empirical realism. Also, the relationship between the ultimate consensus and our factual consensus must be seen in this contrast, which leads us to consider truth to be a “regulative” idea.

The Problem of Truth as a Regulative Idea

Incidentally, Kant states that transcendental idealism or empirical realism is contrasted with transcendental realism or empirical idealism. The latter combination represents the position that “after wrongly supposing

⁴⁵ Other than this part, we can find the notions of this dual structure for instance in A28/B44, A35ff./B52ff., and the references to empirical reality in A582/B610 and A675/B703. Here I do not address his somewhat different approach to the same point in the *Prolegomena*. Citations from Kant are based on the following source and are referred to with the edition (A, B) and page numbers: Immanuel Kant, *Critique of Pure Reason*, trans. Norman Kemp Smith (Basingstoke: Palgrave Macmillan, 2007).

⁴⁶ A369, A370.

that objects of the senses, if they are to be external, must have an existence by themselves, and independently of the senses... finds that, judged from this point of view, all our sensuous representations are inadequate to establish their reality."⁴⁷ In this sense, we can contrast the fallibilism based on metaphysical realism, which is, so to say, the expanded version of Wellmer's criticism, with transcendental-pragmatic consensus theory.⁴⁸

However, the above contrast is not completely analogical to Kant's case. The present two positions seem to have the following points in common: (i) a consensus in any empirical discourse is fallible and (ii) truth as a metaphysical (or transcendental) instance is regulative for every empirical discourse. Apel states:

First, it [the regulative idea of an ultimate consensus] suggests looking for all possible truth criteria and weighing them against each other in order to reach factual but fallible and hence provisory consensus through the argumentative discourse of a real community of investigators.

Second, equally, the regulative idea of an ultimate consensus suggests that we should look for any counter-arguments in order to question every factual consensus of an existent finite research community, thereby keeping open the way of research toward the ultimate consensus of an indefinite community. Only this final consensus, which could no longer be questioned by argument, can and must be equated with the truth we can strive for.⁴⁹

Therefore, no criteria are sufficient truth conditions in empirical discourse; rather, they are the conditions that are necessary for the pursuit of a regulative ultimate consensus. I will return to this point

⁴⁷ A369.

⁴⁸ We may also be able to count Barry Stroud on Wellmer's side. On the one hand, Stroud criticizes transcendental arguments as implicitly accepting a "verification principle" that denies any reality that is unverifiable for humans; Barry Stroud, *Understanding Human Knowledge* (New York: Oxford University Press, 2000), ch. 2. On the other hand, he defends transcendental realism against Kant's critique; Barry Stroud, *The Significance of Philosophical Scepticism* (New York: Oxford University Press, 1984), ch. 4.

⁴⁹ Apel, "Transcendental Semiotics and Truth," 77.

later. For present purposes, it is especially important to note that, like Habermas, Apel seems, after all, to “undermine” the distinction between the meaning and criteria of truth. His account is apparently limited to providing an explication of the meaning of truth but does not present a criterion for truth, because the idea of ultimate consensus does not permit us to know whether our empirical consensus (i.e., about each proposition) is true or false. Wellmer’s criticism looks plausible in this context.

Nonetheless, Apel emphasizes the regulative character of the ultimate consensus, which should lead to our continuation of discursive arguments. Therefore, the key to the problem should involve examination of how this regulative idea can help us provide a criteriologically relevant account of truth.

Interestingly enough, Popper⁵⁰ makes some similar remarks about the role of truth: “it [metaphysical realism] forms a kind of background that gives point to our search for truth. Rational discussion, that is, critical argument in the interest of getting nearer to the truth, would be pointless without an objective reality.”⁵¹ This similarity is quite understandable if we focus on the fact that both Apel and Popper consider empirical discourse to be fallible. To proceed meaningfully with any investigation, despite the fallibility of every factual conclusion, we need an idea that points us in the right direction. How, then, should we understand the difference in the meaning of truth, namely the independent objectivity and the ultimate consensus?

It should be noted here that the concept of the ultimate consensus is not *the absolute* that is supposed by Wellmer and, hence, not an “unknowable (in the long run).”⁵² Kant considers regulative ideas (namely the soul, the world, and God) to be completely transcendent over knowledge.⁵³ In contrast, Peirce and Apel understand it as “the

⁵⁰ See also n.30.

⁵¹ Popper, *Realism and the Aim of Science*, 81.

⁵² Apel, “Pragmatism as Sense-Critical Realism,” 469.

⁵³ They are not “constitutive principles for the extension of our knowledge to more objects than experience can give,” but “regulative principles of the systematic unity of the manifold of empirical knowledge in general” (A671/B699).

knowable (in the long run).⁵⁴ Thus, it is reasonable to suppose that the idea of the ultimate consensus is not merely regulative in Kant's sense. Apel points out that Peirce, in his semiotic interpretation of Kant, has put the latter's regulative principle of experience in the place of Kant's constitutive principles of experience, "on the assumption that the regulative principles in the long run turn out to be constitutive."⁵⁵

I consider that the same applies to Apel's theory of truth. In particular, Apel asserts that the intersubjective validity of transcendental-pragmatic insights into the presuppositions of argumentative discourse does indeed require consensus, but that these insights are a priori "capable of consensus" (*konsensfähig*).⁵⁶ Such insights are explicable in the form of "a sentence which one cannot understand without knowing that it is true"; for instance that "I argue, therefore I recognize the rules—including also the ethical norms—of an unlimited ideal communication community."⁵⁷ Thus, even in a factual discourse, we have some foresight regarding the promised ultimate consensus.

Metaphysical realism cannot propose such an account of truth. Truth may indeed serve as a "background" or a regulative idea for our empirical discourses, but strictly speaking, it cannot serve as a directive of the discourses because, as Wellmer remarks, it is separated from our historical world. Hence, a proponent of this position must sooner or later choose whether to proceed with an investigation (as Popper does), just believing without any good reason that it is in the right direction, or simply abandon the regulative idea entirely and retreat to a far weaker concept of truth, which is the move that Wellmer in actuality makes. If neither of these options is satisfactory, then we should attach greater importance to the transcendental-pragmatic consensus theory of truth.

⁵⁴ Apel, "Pragmatism as Sense-Critical Realism," 469.

⁵⁵ Apel, *Towards a Transformation of Philosophy*, 88.

⁵⁶ Karl-Otto Apel (Ed.), "Fallibilismus, Konsensstheorie der Wahrheit und Letztbegründung," in his *Auseinandersetzungen in Erprobung des transzendentalpragmatischen Ansatzes* (Frankfurt: Suhrkamp, 1998), 186.

⁵⁷ Apel, "Fallibilismus, Konsensstheorie der Wahrheit und Letztbegründung," 185.

Ultimate Grounding and Conditions for Truth

Thus far, I have offered a consistently favorable illustration of the transcendental-pragmatic consensus theory of truth. However, I still find some problems with this theory. Among them are: (i) Putnam's criticism of Apel concerning *sufficient* and *necessary* conditions for true beliefs, and (ii) the problem of the meaning of "ultimate" grounding, which must not be infallible, but is nevertheless performed in factual discourses, which Apel himself regarded as fallible. In my opinion, these two problems are related to each other.

Putnam points out that transcendental pragmatics, especially in regard to its grounding of discourse ethics, has not fully addressed that for which it is accountable. That is, the justification for "the norms and maxims of discourse ethics" that will "be accepted by all the participants" of an ideal discussion is not "that they are the outcome of an indefinitely prolonged Peircean inquiry at all."⁵⁸ What Apel should show here is not only that norms and maxims are *necessary* conditions for truth but that they are *sufficient* conditions for it as well.⁵⁹ These criticisms are reasonable. Apel reduces the criteriological relevance of his account of truth to specifying the necessary conditions for truth rather than the sufficient ones.

On the other hand, Apel explains transcendental-pragmatic insights regarding the presuppositions of argumentative discourse by claiming that they do require consensus but that they are, nevertheless, a priori capable of consensus. This requires an explanation. Indeed, neither of his arguments is satisfactory. How, then, can we refine them?

First, we should redefine the term "truth criterion" employed by Habermas and Apel by means of conditions, namely, sufficient and necessary conditions for truth. Generally speaking, a criterion for truth is that which allows us to identify something as true. We cannot name something "true" if it only fulfills the conditions necessary for truth. In contrast, it will be true if it fulfills only one of the conditions sufficient

⁵⁸ Putnam, *Collapse of the Fact/Value Dichotomy*, 125.

⁵⁹ Putnam, *Collapse of the Fact/Value Dichotomy*, 125–6.

for truth. Apel's appeal to the regulative idea seems to cover only the first point, which is why Putnam criticizes him. Thus, it must be shown here that the transcendental-pragmatic account does or can contain the notion of sufficient conditions for truth.

As I see it, we do not have to assign necessary and sufficient conditions for truth to the same dimension of our argumentative discourses and consensuses, that is, to the transcendental/ideal realm. It is obvious that an argument's being ultimately agreed upon in the ideal discourse is a necessary and sufficient condition for truth. However, this does not imply that there cannot be other sufficient conditions. In my opinion, we can consider ultimate grounding in an empirical discourse⁶⁰ as an economical sufficient condition. This avoids the difficulty that an argument may not be agreed upon in an unlimited ideal discourse even though it is ultimately grounded in an empirical discourse. This point is a great advantage of the transcendental-pragmatic account of ultimate grounding, as we can, in actuality, attain a truth. This account can address not only the meaning of the phrase "a priori capable of consensus" but also Putnam's criticisms. We now have a sufficient condition for truth; norms and maxims of the ideal discourse will arise not from the ideal discourse itself as its conclusion but from empirical ultimate groundings that are sufficient to fulfill the necessary condition for truth, that is, from a prioris to be agreed upon in the ideal discourse.

However, I should emphasize that ultimate grounding is restricted to arguments that recursively or self-referentially state the presuppositions of argumentative discourses (e.g., the assertion, "I recognize you as an equal partner of an argumentative discourse," which is uncircumventable because its negation results in a performative self-contradiction). Thus, almost all propositions, arguments, and theories remain open to Peircean inquiry and, in this sense, are fallible.

In contrast, fallibilism based on metaphysical realism seems to reveal its limits. We can safely say that being correspondent to things in

⁶⁰ Apel's most faithful supporter, Wolfgang Kuhlmann, emphasizes that this grounding itself is not an outcome in the long run of the unlimited ideal discourse, but must be attained "here and now"; Wolfgang Kuhlmann, *Unhintergebarkeit* (Würzburg: Königshausen & Neumann, 2009), 51.

themselves is a sufficient condition for truth, while surviving each critical trial for falsification is a necessary condition.⁶¹ However, in that case, we can never practically satisfy a sufficient condition, and contrary to Popper's view, we can be sure that we are approaching truth only if we already know that the number of necessary conditions, that is, the number of critical trials, is finite and can, in principle, be exhausted. But this assumption is of course contradictory to the Popperian idea of infinite critical trials.

Finally, I would like to enumerate the consequences of the transcendental-pragmatic consensus theory of truth. (i) This consensus theory illuminates the possibility of an alternative to metaphysical-realistic accounts of truth through its transcendental anti-realistic character. We can only really approximate truth through empirical discourses, which is what Popper failed to show with his fallibilism based on metaphysical realism. (ii) Through this consensus theory, we can attain a criteriologically relevant account of truth. This account not only integrates all traditional—and individually unsatisfactory—truth criteria but also (in respect of the task of philosophy) has an advantage over the mere formal explications of the meaning offered by Tarski, by (in a broad sense) the deflationists, and by modest hermeneutic accounts such as those provided by Wellmer. On the one hand, it explicates the meaning of truth as being ultimately agreed upon in the unlimited ideal communication community. On the other hand, it can explain the criterion for truth in terms of necessary and sufficient conditions for truth. The necessary condition is the ultimate consensus, and the sufficient one is the ultimate grounding in empirical discourses. Once a validity claim concerning the norm of the argumentative discourse itself is redeemed in an empirical argumentative discourse, then it is a priori capable of ultimate consensus.⁶²

⁶¹ This understanding reflects, in particular, Popper's thoughts on the philosophy of science.

⁶² This study was supported by JSPS KAKENHI no. 26-5156. I would like to express my gratitude to the editors of this book for reading the previous version and providing very helpful comments.

Michihito Yoshime is currently Research Fellow of the Japan Society for the Promotion of Science and a faculty member at the Faculty of Humanities at Sophia University. His recent publications include the article “What Is the Unlimited Communication Community? Transcendental Pragmatics as Contemporary Fichteanism,” in *Fichte and Transcendental Philosophy* (Palgrave Macmillan, 2014). Among his main research interests are Kant, German idealism, transcendental pragmatics, and discourse ethics.

13

On Jürgen Habermas's Cognitive Theory of Morality

Yasuyuki Funaba

Jürgen Habermas differentiates assertoric sentences used in constative speech acts from normative sentences used in regulative speech acts. Following this differentiation, I first would like to argue that in the same way as normative sentences constitute rightness, assertoric sentences constitute the truth, although both types of sentences can be used cognitively in different ways. Next I examine Hilary Putnam's criticism of Habermas, in order to gain a better understanding of his position. It will become obvious that no moral norms can be grounded in an absolute way (*letztbegründet*), if one assumes a meta-ethical-cognitivist moral theory without the presupposition of a moral realism; instead, one can only say that for some moral norms there are just no alternatives (*alternativenlos*). And this lack of alternatives (*Alternativenlosigkeit*) manifests itself again and again in all moral argumentation. However, there is

Y. Funaba (✉)

Department of Philosophy and History of Philosophy, Graduate School of Letters, Osaka University, Osaka, Japan
e-mail: funacho@let.osaka-u.ac.jp

© The Author(s) 2016

H. Kim, S. Hoeltzel (eds.), *Transcendental Inquiry*,
DOI 10.1007/978-3-319-40715-9_13

285

no other possible conclusion, as—according to the discourse theory—it is precisely through argumentation that it becomes clear whether a norm is right or not. Even if Putnam criticizes Habermas’s theory with the word “minimalism,” this shows that he understands the full spectrum of the theory that Habermas proposes.¹

Constative Statements Versus Normative Statements

Habermas describes his moral theory as a cognitivist theory.² It is cognitivist for the reason that Habermas understands “normative rightness . . . as a validity claim that is analogous to truth.”³ In meta-ethics, the difference between cognitivism and noncognitivism lies in whether the claim of moral rightness can be identified in a cognitive manner. However, both positions share the common view that we can gain “scientific findings only from the objects of the spatial physical world.”⁴ For this reason, emotivists like Charles L. Stevenson, who follows up on the premise of logical positivism and ascribes all moral statements to subjective feelings, represent a noncognitivist approach, whereas naturalists like David K. Lewis represent the position of meta-ethical cognitivism, because they derive “the meaning of the moral predicates” “from the predicates of an empirical description of the world of physical objects.”⁵ In meta-ethics, the difference between cognitivism and noncognitivism is dependent upon whether moral findings can be derived from physical objects, and if one thinks that they can be derived from these objects, then one also thinks that the

¹ In short, Putnam argues for a cognitive theory of morality that presupposes a moral realism, while Habermas defends a cognitive theory of morality without presupposing moral realism. That is why Habermas cannot help being a moral minimalist, but that does not make him wrong.

² See EA 11. In this essay I use the expression “moral theory,” because Habermas himself said that the name of “discourse ethics” might have been misleading (see EA 101).

³ EA 11.

⁴ See Matthias Lutz-Bachmann, *Ethik* (Stuttgart: Reclam, 2013), 29–30.

⁵ Lutz-Bachmann, *Ethik*, 33.

truth and the rightness of moral statements can be ascertained. Under the premise of the existence of moral objects, meta-ethical cognitivism is finally confirmed, and therefore the truth of the moral statements presumes a moral—although not always just a naturalistic—realism. Habermas, however, does not follow this way of thinking. He thinks that the validity claim that is analogous to truth is possible but does not presuppose the existence of moral objects. Without assuming a moral realism, consequently, Habermas represents the school of cognitivism and is of the view that the rightness of moral statements can be determined.

In his work *Moralbewusstsein und kommunikatives Handeln* (Moral Consciousness and Communicative Action), Habermas differentiates the constative statement from the normative: “norms [are] dependent on the continuous reproduction of legitimately ordered interpersonal relationships. . . . In contrast to this, we need to assume the concept that things exist, regardless of whether they can be stated through true sentences or not.”⁶ Norms have—as long as they are social norms—a social validity, when they are intersubjectively recognized as such in a given society. However, we can differentiate here between social validity (*Geltung*) and normative validity (*Gültigkeit*).⁷ With Kant’s statement, “if he has committed murder, he must die,”⁸ it is possible, based on the acknowledged social validity of norms in a given society, to claim normative rightness. Kant’s statement is possible because he approves the death penalty, which gives this statement its rightfulness. Of course it is also possible to question the death penalty itself. This means the social validity (*soziale Geltung*) of the death penalty itself, so it is

⁶ “Normen [sind] darauf angewiesen, dass legitim geordnete interpersonale Beziehungen immer wieder hergestellt werden. . . . Demgegenüber sind wir konzeptuell zu der Annahme genötigt, dass Sachverhalte auch unabhängig davon existieren, ob sie mit Hilfe wahrer Sätze konstatiert werden oder nicht” (MkH 71). Without this differentiation, naturalistic cognitivism revokes the difference between empirically “true” and morally “right”; see Lutz-Bachmann, *Ethik*, 33.

⁷ Habermas discusses this in the context of the “ambiguous nature of ought-validity” (*Sollgeltung*); see MkH 71.

⁸ Immanuel Kant, *Die Metaphysik der Sitten*, in *Kant's gesammelte Schriften*, ed. Royal Prussian (later German and Berlin-Brandenburg) Academy of Sciences, 29 vols (Berlin: George Reimer (later Walter de Gruyter), 1900–), vol. 6, p. 333.

questioned whether the social validity of norms is valid (*gültig*). In this case, the point is not whether the social validity (*Geltung*) of the norm is an actual fact in a society, but whether this norm is valid (*gültig*). In other words, if one follows the existing norm and says, “if he has committed murder, he must die,” without questioning the social validity of this statement, then one claims not only the rightness of the statement, but also the rightness of the norm itself. If the normative rightness is claimed with a normative statement, then at the same time the normative rightness of the norm that justifies the normative statement is claimed. Through this process, the validity of the norms is continually reconfirmed. Consequently, if there were not any normative statements, there would not be any validity of norms. With the acknowledgment of the rightness of each normative statement, the rightness of the norms themselves is acknowledged. And this is what constitutes the basic difference between the constative and the normative statement.

“Normative validity claims convey obviously a reciprocal dependence between the language and the social world that does not exist for the relation between language and the objective world.”⁹ Unlike in the case of normative statements and norms, the actual state of affairs exists regardless of the constative statements. Whether a fact exists or not does not depend on a speech act like “the snow is white.” Habermas gives the statement “iron is magnetic” as an example of a constative statement.¹⁰ Iron is magnetic, but not because the statement “iron is magnetic” has been made and its claim to truth has been intersubjectively acknowledged. The fact that iron is magnetic exists regardless of the validity of the statement “iron is magnetic.” While norms can be changed through normative statements, no facts change merely because a constative statement has been made. Whether the claim to truth that has been raised with the constative statement can be intersubjectively acknowledged or not is dependent upon whether the fact is described appropriately by the respective statement. Constative statements become true only when they linguistically express exactly the facts that exist

⁹ Mkh 71.

¹⁰ Mkh 70.

regardless of the statements. Wanting to know “the conditions for the validity of moral judgments” suggests directly “the transition to a logic of practical discourses,” while wanting to know “the conditions for the validity of empirical judgments requires epistemological and scientific-philosophical considerations that are initially independent of a logic of theoretical discourses.”¹¹ Through the process of theoretical discourses alone, constative statements are not necessarily made true. In contrast, normative statements are only right when their claim to rightness is intersubjectively acknowledged through the process of practical discourses; they cannot be valid without this process.

The Lack of Alternatives for the Presuppositions and Their Justification

So what are the “conditions for the validity of moral judgments” and consequently the conditions for the validity of moral norms? According to Habermas, the validity of moral norms is put into question when we somehow lose a mutual understanding regarding these moral norms. For example, a certain person considers the moral norm M1 as valid, while another person considers M2 as valid. In this case, it becomes clear whether M1 or M2 (or possibly a third option) is valid only when both persons present their point of view in moral argumentation with appropriate arguments and come to a consensus on the valid norm after the discussion. As the lost mutual understanding regarding the valid moral norm is regained precisely through moral argumentation, it becomes clear that the principle of universalization (U) is the moral principle that regulates moral argumentation. This (U), which says that “all affected can accept the consequences and the side effects that [the norm’s] general observance can be anticipated to have for the satisfaction of everyone’s interests, and the consequences are preferred to those of known alternative possibilities for regulation,” is the condition that every valid norm has to fulfill. Considering the interests of all affected,

¹¹ M_kH 72.

it is possible to reach through such argumentation a consensus on the validity of norms. Based on this consensus, the affected acknowledge the validity of the moral norm that follows (U). This is precisely the point of consensual cognitivism.

So how can the validity of (U) be justified as a moral principle? According to Habermas, the attempt to solve normative problems through argumentation is a “practice of everyday communication” (*kommunikative Alltagspraxis*).¹² Does this mean that the moral principle is only reflexively derived from what we do in our everyday life? If the rightness of the moral principle were based on the everyday facts, there would be no difference between the constative statement and the normative statement, because the validity of every moral norm rests on the validity of the moral principle, and the validity of every constative statement is based on the actual facts. So why is a normative statement right? Because it is derived from the right moral principle. And why is the moral principle right? Because it is an exact reconstruction of the existing facts. This explains the rightness of a normative statement. However, by reflecting on the practices of everyday life, Habermas tries not only to explain the moral principle, but also to justify it. However, Habermas does not say that (U) is valid as a moral principle because it reconstructs exactly what exists in the practices of everyday life. As justification, Habermas assumes the “examination of the pragmatic conditions for the argumentation in general”¹³ and derives “from their propositional content [the] principle of universalisation (U).”¹⁴

Drawing on Robert Alexy’s discussion,¹⁵ Habermas states that the rules of the discourse are “inescapable presuppositions” that determine “that nobody who could make a relevant contribution [to the discussion] may be excluded.” They grant “all participants . . . an equal opportunity to make contributions” and demand conditions of communication that enable each

¹² Mkh 77.

¹³ Mkh 86.

¹⁴ Mkh 93.

¹⁵ Robert Alexy, “Eine Theorie des praktischen Diskurses,” in *Normenbegründung, Normendurchsetzung*, ed. Willi Oelmüller (Paderborn: Schöningh, 1978), 40–1.

and every person to exercise the “right to universal access to and equal participation in” the discourse: “The presuppositions themselves can be identified by showing one, who at first denies the hypothetically offered reconstructions, how he gets entangled in performative contradictions.”¹⁶ However, this only shows the “lack of alternatives for these rules in the practical argumentation,”¹⁷ but it does not provide a justification for the lack of alternatives (*Alternativenlosigkeit*) itself. So what does it mean that these rules lack alternatives, and that at the same time the lack of alternatives cannot be justified?

If the lack of alternatives for these rules cannot be justified, then these rules prove to be without any alternative every time one shows that he or she who opposes the validity of these rules is committing a performative contradiction. That he or she who challenges the validity of these rules has to follow the challenged rules in order to do so can be explained by pointing out the performative contradiction, which at the same time gives an explanation for the lack of alternatives. It is important to point out the performative contradiction, and thus to explain the lack of alternatives for the rules, every time the opponent speaks out against the validity of the rules. If the lack of alternatives could be justified even once, it would be (at least in principle) not necessary to point out the performative contradiction every single time the opponent speaks out against the validity of the rules. This would be because, even without pointing out the performative contradiction to the opponent, it would be clear that he or she ought to follow those rules, because they are in principle justified. Habermas contradicts Apel, who postulates the “ultimate justification” (*Letztbegründung*),¹⁸ and he thinks that the lack

¹⁶ “Die Präsuppositionen selbst können nun in der Weise identifiziert werden, dass man demjenigen, der die zunächst hypothetisch angebotenen Rekonstruktionen bestreitet, vor Augen führt, wie er sich in performative Widersprüche verwickelt” (MkH 100). Cf. Apel: “By a transcendental pragmatic self-contradiction, I understand a performative contradiction between the content of a proposition and the intentional content of the act of proposing the proposition”; Karl-Otto Apel, *Understanding and Explanation: A Transcendental-Pragmatic Perspective* (Cambridge, MA: MIT Press, 1985), 8.

¹⁷ MkH 105.

¹⁸ See MkH 106.

of alternatives and the inescapability of the presuppositions of the argumentation do not exist independently of the act of argumentation that is exercised against the opponent every time in order to show him or her the performative contradiction. On the contrary, they would get confirmed every time such an act of argumentation is performed.

The situation in which participants in the argumentation have to assume these presuppositions inescapably is precisely the situation in which (U) is valid for the moral argumentation: “from the mentioned discourse rules it becomes clear that a disputed norm will only find approval amongst the participants of a practical discourse, if (U) is valid.”¹⁹ So in this situation it will be discussed whether (U) is valid as moral principle: the validity of (U) as moral principle can always be justified at the same time as the validity of the presuppositions as rules is confirmed by the argumentation by pointing out the performative contradiction to the opponent who objects to the validity of these rules. With the argumentation, the validity of the presupposition and of (U) is also latently confirmed, if one argues without objecting to the validity of the presuppositions. The validity of the presuppositions that make the argumentation possible and of (U), which regulates the moral argumentation as a moral principle, is also acknowledged every time the validity of each single moral norm is acknowledged in the moral argumentation by the participants through consensus.

Putnam’s Criticism of Consensual Cognitivism

At this point I would like to discuss Putnam’s criticism of Habermas, in order to clarify further Habermas’s theory about normative statements.²⁰ Putnam criticises Habermas in two different essays, but the core of

¹⁹ Mkh 103.

²⁰ According to Habermas, we can say as the result of our consensus whether a norm is right or not. This is the point of consensual cognitivism.

the criticism is the same in both cases. Putnam discusses two possible opinions which he believes come from Habermas.²¹

The first possibility is the assumed opinion derived from the discussion that was developed by Apel based on transcendental pragmatics.²² According to Putnam, “the heart of Apel’s position is that, following Peirce, it identifies truth with *what would be agreed upon in the limit of indefinitely continued discussion* . . . Apel’s important move is to apply this identification also to *ethical claims*, indeed to *all* discourse.”²³ According to Putnam, Apel is of the view that the right moral statement is the result of ideal argumentation. Consequently, it is not the case that we reach a consensus through ideal argumentation because the moral statement in question is right; on the contrary, the statement is right because we reach a consensus through ideal argumentation. According to transcendental pragmatics, one follows (U) and the rules of argumentation by deriving individual moral norms, where the principle and the rules are the transcendental conditions for the derivation of the individual moral norms.²⁴ If Apel, as quoted above, takes an important step forward, the principle and the rules show “not only an ideal condition for rational understanding [*Verständigung*], but in reality at the same time a condition for ideal mutual understanding [*Verständigtsein*],” as Albrecht Wellmer says.²⁵ With the following example, Putnam argues critically that neither (U) nor the rules of argumentation are a “sufficient condition”²⁶ for gaining the right moral norms: “a father

²¹ Putnam says that “very likely Habermas’s actual response will turn out to be different from both”; Hilary Putnam, “Values and Norms,” in *The Collapse of the Fact/Value Dichotomy and Other Essays* (Cambridge, MA: Harvard University Press, 2002), 120. According to my understanding, Habermas would not hold the first view, and the second could not be criticized.

²² The discussion developed in two essays that were published before *Theorie des kommunikativen Handelns*: “Wahrheitstheorien” (1972), and “Was heißt Universalpragmatik” (1976), in Habermas, *Vorstudien zur Theorie des kommunikativen Handelns* (Frankfurt: Suhrkamp, 1984), 127–83 and 353–440.

²³ Putnam, “Values and Norms,” 122.

²⁴ See Wolfgang Kuhlmann, *Reflexive Letztbegründung* (Freiburg: Alber, 1985), 73.

²⁵ Albrecht Wellmer, *Ethik und Dialog* (Frankfurt: Suhrkamp, 1986), 101–2.

²⁶ Putnam, “Values and Norms,” 126.

engages in psychological cruelty by teasing his child, while denying (either because he is obtuse or because of a streak of sadism) that the child's tears are really 'serious.' 'He has to learn to take it,' the father says."²⁷ According to Putnam's criticism, there is a strong possibility that the community cannot agree on regarding the father's action as abuse, even if the members of the community wholeheartedly want to do the right thing and have a liking for rational argumentation, when there are persons that are as intellectually restricted as the father. In this case, the community would not necessarily reach a consensus on the right answer, even if the circumstances for the conversation were ideal and the members discussed the case for long enough. According to Putnam's criticism, to identify the father's action as abuse, one needs the moral capability to realize "that someone is, for example, 'suffering unnecessarily' as opposed to 'learning to take it,'" and this capability "is interwoven with our . . . mastery of moral vocabulary itself."²⁸ Without presupposing this moral capability as a condition for ideal argumentation, we cannot reach a consensus on the right moral norms, even when following (U) and the rules of argumentation. But if this capability is necessary, the consensus does not carry much importance for the rightness of the moral norms. A moral norm is not right just because we have agreed on it, but because persons equipped with the necessary moral capabilities judge that this moral norm is right. Without the premise of the particular moral capability, the rules of argumentation are finally no sufficient condition for gaining the right moral norms. With the premise of this moral capability, on the other hand, the rules of argumentation are not necessary for establishing the right moral norms. These unwanted conclusions can be drawn because one wants to see the rules of argumentation as sufficient conditions for the right moral norms. However, Habermas does not draw these

²⁷ Putnam, "Values and Norms," 127. Cf. Putnam, "Antwort auf Jürgen Habermas," in *Hilary Putnam und die Tradition des Pragmatismus*, ed. Marie-Luise Raters and Marcus Willaschek (Frankfurt: Suhrkamp, 2002), 307.

²⁸ Putnam, "Values and Norms," 128.

conclusions, because, as pointed out earlier, he demands that in every argumentation the validity of the rules of argumentation itself be questioned again and again.

The second possible opinion concerns where Putnam questions the fact/value dichotomy and cannot factor terms like “kind” or “cruel” into two components, that is, a descriptive and an evaluative one: “the fact that an act is cruel or kind... is available only through the lens of value concepts.”²⁹ Putnam also objects to the norm/value dichotomy, because he thinks that there are no norms that can be expressed without using words expressing values. For this reason and on the grounds of his value realism, he represents the school of cognitivism with regard to moral norms. This means that the rightness of moral norms can be identified under the condition that moral objects exist. It goes without saying that Habermas cannot accept this alternative, because, as pointed out above, he differentiates between the constative statement and the normative statement. “Correct moral judgments owe their universal validity not to their corroboration by the objective world like true empirical judgments,”³⁰ he writes. However, he also expresses a view very different from that of Apel: “it is by no means self-evident that rules that are inescapable within discourse can claim to be also valid for regulating actions outside of discourses.”³¹ The rules of argumentation regulate merely the process for “proving the validity of suggested and hypothetically assumed norms,”³² but they do not provide us with concrete norms of action that are also valid outside of discourses. This point of view is nothing other than the second possible opinion assumed by Putnam, which represents a “more ‘minimalist’ position.”³³ If Habermas—as a moral minimalist—has

²⁹ Putnam, “Values and Norms,” 119.

³⁰ Habermas, “Werte und Normen: Ein Kommentar zu Hilary Putnams Kantischem Pragmatismus,” in *Hilary Putnam und die Tradition des Pragmatismus*, 299.

³¹ Mkh 96.

³² Mkh 113.

³³ Putnam, “Values and Norms,” 122.

to reduce statements about moral norms “strictly” to the rules of argumentation, then Putnam criticizes him on the grounds that he can only say, “discuss according to the rules of argumentation the question whether you are allowed to be cruel or not!” However, he cannot say, “don’t be so cruel!”³⁴ Putnam complains that Habermas did not reply to this criticism,³⁵ but in my opinion, Habermas did not need to reply to it.

In the debate with John Rawls, Habermas once argued critically that the principles of a fair society in the so-called original position are not chosen from the participants’ perspective, but only from the perspective of the theorist himself. He also pointed out that Rawls’s theory involves the risk of ending up in “philosophical paternalism,”³⁶ in which the philosopher alone draws the conclusion from the discussion. It goes without saying that Habermas could try—as a participant in moral argumentation—to discuss the rightness of the norm “don’t be so cruel!” and to find a relevant consensus with the other participants through argumentation. However, in contrast to Putnam, Habermas of course does not present this norm from his point of view as moral philosopher.³⁷ As Putnam says, “rather than undertake the task of producing a ‘final’ ethical system, a final set of rules of conduct, what Habermas offers us instead is a rule for how to conduct our inevitable disagreements over the first-order rules that should govern our conduct.”³⁸ Putnam’s understanding of Habermas on this point is by no means wrong. If the philosopher as a result of his moral philosophy just offers us—without proper moral argumentation—the norms that claim to be valid regardless of a consensus, then he acts inconsequently as he postulates the consensual cognitivism for the individual moral norm and at the same time for the rules of argumentation.

³⁴ See Putnam, “Antwort auf Jürgen Habermas,” 309.

³⁵ Putnam, “Antwort auf Jürgen Habermas,” 309.

³⁶ EA 119.

³⁷ “The moral theorist can take part as affected person, if applicable as expert, but he cannot direct these discourses” (MkH 104).

³⁸ Putnam, “Values and Norms,” 116.

Conclusion

Moral realism focuses on the discussion that presupposes moral objects which exist regardless of our argumentation. Habermas does not follow this realism, because he explicitly differentiates between the constative statement and the normative statement. He nevertheless constitutes his moral theory in a cognitivist way. And this is nothing other than the argumentative consensus. Valid moral norms become cognitively acknowledged through the consensus that is enabled by argumentation. Consequently, the validity of the moral principle and of the rules of argumentation is acknowledged every time the validity of the individual moral norm is acknowledged.

Yasuyuki Funaba is Associate Professor of Philosophy at the Faculty of Letters of Osaka University. His recent publications include an anthology (with Y. Kato), *Kant and Contemporary Philosophy* [in Japanese] (Kōyōshobō, 2015), as well as a Japanese translation (with Y. Mikoshihira) of the anthology *Recht auf Menschenrechte* (Osaka University Press, 2015). Among his research interests are Kant and Habermas.

Index

A

Alexy, Robert, 290
Anscombe, G. E. M., 187, 245
Anti-realism, 264–265, 273–276
Apel, Karl-Otto
 contrasted with Habermas, 278
 founder of discourse
 ethics, 210–212, 227, 230,
 264, 280
 grounding of discourse ethics, 280
 method of rationally
 definitive justification
 (*Letztbegründung*), 211
 sense-criticism (*Sinnkritik*), 211
 theory of truth, 210, 223,
 264–270, 274–276,
 279–280, 282
 transcendental pragmatics of
 communication, 265, 274,
 280, 293

transcendental reflection, 210,
212–213, 215, 217–218, 224

See also Discourse, argumentative;
Ideal communication
community

Argumentation, rules of,
see Discourse, argumentative

Aristotle, 23–24, 169, 266

Austin, J. L., 210, 234, 243, 268

B

Bacon, Francis, 98
Baggesen, Jens Immanuel, 35
Bardili, Christoph Gottfried, 100
Baumgartner, Alexander, 3
Being-in-the-world, 110, 174
Belnap, Nuel, 246
Berkeley, George, 5, 25, 32, 161
Blumenbach, Johann Friedrich, 90

Brandom, Robert, 241

Burnyeat, Myles, 25

C

Cassirer, Ernst, 169

Davos debate with Heidegger, 169

Cicero, 98

Circle, 65, 83, 99, 125–126, 138

Cohen, Hermann

Continuity, 147

critical idealism, 24, 127–128,

132, 142–143, 152,

154, 156

determination, 145–150, 155, 169

experience, 128–134, 136–138,

141, 153–154

fact (*Faktum*), 129–130, 133–134,

137–141, 144, 153

humanity, 155–156

hypothesis, 127–128, 143–145,

150, 154, 155, 156, 190

judgment, 144–151

origin, 127–130, 143–152, 272

pure intuition, 11, 131

quid juris, 31, 136

sciences (*Wissenschaften*), 129,

130, 133, 135–137,

143–144, 146, 152,

153, 155–156

space and time, 130, 132, 135

task (*Aufgabe*), 130, 136–137,

154, 156

the transcendental, 127–132,

138, 145, 154–155, 156

transcendental method, 127, 129,

136, 138

transcendental

philosophy, 128–129

Collingwood, R. G., 253

Consensus, transcendental

implications of, 272–273

Constructivism

in ancient mathematics, 27–28,

31–32

in Kant, 31, 33–34

in post-Kantian idealism, 82

vs. Berkeleyan idealism, 32

Contradiction

logical, 226, 228, 234–235

vs. nonconformity, 237–238

performative self-

contradiction, 210–211, 213,

227–228, 231, 264, 281 (*see*

also Transcendental argument

forms, performative or

retorsive)

pragmatic, 234–250, 246, 253,

258, 260

question-answer, 233–234, 238,

243–245

mixed, 241

pure, 238–239, 241

semantic, 234–236, 253, 254–255

D

Davidson, Donald, 242, 259

truth-conditional semantics, 210

Descartes, René, 22, 99, 169, 267

Dilthey, Wilhelm, 133, 161, 163, 165

Discourse

argumentative, 210–230, 264,

269, 277, 279–282

rules of, 290, 292, 294–296

uncircumventable conditions

of, 281

- rules of, 227, 290, 292
 - justification for, 214, 280, 291
 - (*see also* Discourse ethics)
- Discourse ethics
 - ultimate grounding of, 264–265, 280–282
 - See also* Apel, Habermas, Putnam, Wellmer
- Dogmatism, 61, 64
- E**
- Emotivism, 286
- Ethics of belief
 - the agnostic imperative, 73–74
 - evidentialism, 72–74, 80
 - fideism, 57, 81
 - nonevidentialism, 57, 81–82
 - practical (Kant), 46–57, 75
 - pure-rational (Fichte), 57, 82
- Existence-philosophy, 161
- F**
- Fact/value distinction, 273, 293, 295
 - See also* Norm/value distinction
- Fallibilism, 263, 273, 277, 281–282
- Fichte, Johann Gottlieb
 - absolute self-sufficiency, 57
 - being *vs.* acting, 37–39, 42, 49, 63, 67
 - cognitive constructivism, 27–28, 31
 - conditions for the possibility of
 - experience, 76, 153
 - developmental psychology, 89
 - facts of consciousness, 89–91, 93–94, 103–104
 - faith or belief (*Glaube*), 57, 81, 98
 - freedom, 35–40, 35–54, 92, 101, 103
 - formal, 38–40, 41, 44, 49, 50
 - material, 38–40, 43, 44, 49, 51–52
 - moral, 43
 - system of freedom, 35, 36
 - of thinking, 37
- immaterialism, 57
- intellectual intuition, 51
- Kant, 35–37, 42
 - criticisms of, 75, 96, 108
 - influences from, 35
- morality, 41–43, 52, 54, 103
 - moral freedom, 43
 - moral law, 42, 46, 47–49, 51, 53, 102
 - moral world order, 103
 - radical evil, 44–45
- ontotheology, 57, 79
- reason, 46, 68, 75, 77
 - nature of, 36–37, 41, 44–45, 80
 - practical, 52, 72–80
 - rational justification, 56–57
 - spontaneity of, 50–51
- religion, 136, 184
- self-positing, 37–38, 45, 59, 67
- theism, 55, 57, 66, 70, 82
- transcendental epistemology, 56, 69, 72, 79
- transcendental idealism, 32, 57, 65, 83–84, 91, 132–133
- transcendental ontology, 55–56
- transcendental
 - philosophy, 83–104
 - after Jena period, 85
 - idealism *vs.* dogmatism, 61, 64
 - special standpoint of, 60, 62, 64, 84, 97, 112, 117, 123
 - therapeutic dimension of, 98–99

- Finitude, 174
 Fink, Hans, 197
 Finnis, John, 182
 Frege, Gottlob, 25
 Fries, Jakob Friedrich, 93, 131
- H**
 Habermas, Jürgen
 cognitivism without realism, 287, 295
 consensual cognitivism, 290, 292, 296
 consensus theory of truth, 265, 268–270
 contrasted with Apel, 184
 debate with Rawls, 296
 interpretation of Heidegger, 162
 moral minimalism, 286
 principle of universalization (U), 289
 justification of, 290
 Putnam's criticisms, 281
 See also Ideal speech situation
 Hegel, G. W. F.
 the Absolute, 113, 114
 cognition, 113, 114–116
 consciousness
 configuration of
 consciousness, 120, 122
 self-consciousness, 116, 122
 transformation of
 consciousness, 120–124
 empirical objects, 110
 experience, 32, 108–110, 118–121
 effective experience, 118
 metaphysics, 113–115, 121, 124–126
 method
 criterion-less critique, 116–119
 critique of Kant's
 procedure, 108, 113–115
 developmental, 107–126
 dialectical, 121, 123–125
 phenomenological, 108, 115, 117
 rational justification, 122
 self-criticism, 121
 reality, 117, 121
 science, 113, 115, 117, 121, 122
 spirit (*Geist*), 125
 Heidegger, Martin
 ambiguous relationship with
 transcendental
 philosophy, 167–171
 attempts to overcome the
 transcendental
 paradigm, 175
 the history of Being, 175
 the “turning” (*Kebr*), 162
 attempts to rethink
 transcendental
 constitution, 161, 162, 165, 177
 in *Being and Time*, 167
 in terms of the “event” (*Ereignis*), 162
 in terms of fundamental
 ontology, 161
 in terms of world-
 disclosure, 161
 criticisms of Husserl, 166
 criticisms of transcendental
 philosophy, 96, 108
 as ahistorical, 161, 171, 172
 as horizontal, 160, 162, 171

- as reflective-
 - representational, 160, 162
- as subjectivist, 88
- as symptomatic of Western
 - modernity, metaphysics, 54, 82, 104, 167, 169, 170
- early alternatives to transcendental
 - philosophy, 177
 - existence-philosophy, 64, 71, 161, 170, 173
 - hermeneutics, 161, 171
 - life-philosophy, 161, 165, 171
- failure to overcome transcendental
 - philosophy, 159–179
 - on German idealism, 163, 164, 169
- Herbart, Johann Friedrich, 93, 131
- Héring, Jean, 183
- Hermeneutics, 161, 171, 173, 179, 210, 263
- Herz, Markus, 1, 30
- Hobbes, Thomas, 31
- Höffe, Ottfried, 53
- Hölderlin, Friedrich, 54
- Humboldt, Wilhelm, 210
- Hume, David, 23, 25, 30, 88, 110
- Husserl, Edmund
 - criticisms of Heidegger, 166
 - and idealism, 34, 160, 163, 165, 166
 - influence on Heidegger, 183
 - and transcendental philosophy, 160, 165, 168, 178
- I
- Ideal communication
 - community, 265, 269, 271, 274, 279, 282
- Ideal speech situation, 269, 271
- Identity
 - law of, 256, 260
 - types of, 25
- J
- Jacobi, Friedrich Heinrich, 94
- K
- Kant, Immanuel
 - analytic (regressive) *vs.* synthetic (progressive)
 - procedure, 134, 135
 - architectonic, 137
 - categories
 - of modality, 2, 13
 - of relation, 7, 13–14
 - transcendental deduction
 - of, 91, 108, 112, 131
 - cognition, 1, 6–8, 16, 24, 31, 33, 42, 89, 113, 114, 133, 139, 141, 152, 212
 - cognitive constructivism, 27, 28, 31–32
 - conditions for the possibility of
 - experience, 76, 153
 - critical philosophy, 2, 22–23, 29–31, 32, 33, 36, 85, 100, 108
 - deduction, 18, 31, 85, 90, 91, 94, 97, 108, 112, 114
 - empirical objects, 110
 - epigenesis of pure reason, 87
 - experience, 1–19, 50, 87, 112, 120, 128, 129, 133, 153, 159, 165, 279

- Kant, Immanuel (*cont.*)
- freedom, 35–54
 - comparative, 49, 50
 - negative *vs.* positive, 36, 52–53
 - political, 54
 - psychological, 48, 49
 - Third Antinomy, 48, 53
 - Transcendental, 35–54
 - intellectual intuition, 21, 33, 51, 58, 59
 - letter to Markus Herz, 1
 - limits of human knowledge, 2, 17, 18
 - moral theology, 75, 81
 - practical reason, 36, 48, 52–53, 72–82
 - postulates of pure practical reason, 47, 75
 - and practical
 - nonevidentialism, 57
 - primacy of practical reason, 100
 - prohibition against
 - lying, 257–258
 - Refutation of Idealism, 108, 190
 - relation to prior tradition, 23
 - Copernican revolution, 24
 - Leibniz, 23
 - Locke, 98
 - Plato, 23
 - Representationalism, 24, 30
 - Second Analogy, 108
 - synthetic unity of
 - apperception, 114
 - things in themselves, 33, 155, 276
 - double-aspect thesis, 32
 - transcendental arguments, 58–64, 97, 108, 110, 112, 113, 181–207, 233–261
 - transcendental philosophy, 1, 3, 11, 85, 120, 127, 128, 142, 159, 169
 - transcendental reflection, 57, 82, 209–231
 - the transcendental subject, 16, 17, 18, 68, 152, 154, 155, 165, 166, 168, 169
- Kierkegaard, Søren, 183, 207
- Kohlberg, Lawrence, 210
- Korsgaard, Christine, 182, 184–186
- Kuhlmann, Wolfgang, 264, 281, 293
- L
- Lange, Friedrich Albert, 131, 137
- Lange, Johann Joachim, 99
- Leibniz, Gottfried Wilhelm, 23, 49
- Levinas, Emmanuel, 165, 170, 179, 181–206
- Lewis, David K., 286
- Life-philosophy, 161, 171
- Linguistic communication
 - fundamental conditions of, 233, 241, 259
 - confirmation of
 - connection, 241, 260
 - confirmation of sincerity, 239, 260
 - confirmation of
 - understanding, 241
 - mutual responding relations, 244, 245, 259, 260
 - question-answer relations, 233, 234
 - and anaphora, 234, 241, 243, 244, 245

and indexical expressions, 242, 243, 245
 and responding relations, 238, 240, 245, 251, 255, 256, 258, 259, 260
 transcendental conditions
 of, 253, 255
See also Discourse, rules of;
 presuppositions; speech acts
 Lipps, Hans, 183
 Locke, John, 87, 98
 Løgstrup, K. E., 181–184, 196–206
 approach to transcendental
 argumentation, 182, 199, 205

M
 MacIntyre, Alasdair, 107, 197, 198
 Mandeville, Bernard, 186
 Marx, Karl, 34, 186
 Merleau-Ponty, Maurice, 110, 164, 165
 Metaethics, 286
 cognitivism *vs.* noncognitivism, 286–287
 Misch, Georg, 163
 Moore, G. E., 32

N
 Natorp, Paul, 24
 Neo-Kantianism
 anthropological, 93
 influence on Heidegger, 160
See also Cohen; Natorp; Rickert
 Nietzsche, Friedrich, 186
 Norm/value distinction, 295
See also Fact/value distinction

P

Parmenides
 account of cognition, 26
 and German idealism, 26
 and Plato, 25, 28–29
 Peirce, Charles Sanders
 consensus theory of truth, 265, 268, 275
 semiotic interpretation of
 Kant, 265, 279
 Perpich, Diane, 182, 187–206
 Pestalozzi, Johann Heinrich, 89, 100
 Phenomenology, 108–115, 118, 122, 160, 163–166, 167, 168, 172, 178
 early critiques of, 163
 Philosophy, history of, 22–23, 26, 35, 99, 206
 modern attitudes to, 97
 Plato
 Intuition, 29, 33
 Representationalism, 24, 29
 theory of forms, 24, 26, 28, 33
 Popper, Karl, 272, 278, 279, 282
 Presuppositions
 pragmatic, 264, 279, 280
 semantic, 246, 247, 248, 249, 250, 251, 252, 260
 Prichard, H. A., 185, 192
 A priori concepts, 7, 10, 84, 87, 112, 128
 Psychologism, 127
 Psychology, 89, 131, 142, 152, 153, 154, 212
 Puntel, Lorenz B., 267
 Putnam, Hilary
 criticisms of consensus theory
 of truth, 264

Putnam, Hilary (*cont.*)
 criticisms of discourse ethics, 280
 criticisms of Habermas, 285,
 286, 292

Q

Quine, W. V. O., 22

R

Rawls, John, 296

Realism

metaphysical, 25, 26, 27, 28, 33,
 272, 274, 277, 278, 279,
 281, 282

moral, 285, 287, 297

as prephilosophical outlook, 63

sense-critical, 274, 275

in Western epistemology, 25

Reinhold, Karl Leonhard, 43, 58, 100

Rickert, Heinrich, 160

Rubin, Edgar, 183

S

Sartre, Jean-Paul, 35

Scheler, Max, 183

Schelling, F. W. J., 31, 32, 93, 95,
 116, 163, 164, 169, 175

Schmid, C. C. E., 93

Schopenhauer, Arthur, 98

Schulze, Gottlob Ernst, 93

Searle, John

criticism of Heidegger, 166, 173

critique of phenomenology, 164,
 166, 167

intentionalist semantics, 210

Self-contradiction, *see* Contradiction

Sentences

assertoric, 285

normative, 285

Skepticism

cognitive, 27

moral, 182, 187, 204, 205

varieties of, 188

Speech acts

constative, 269, 285

illocutionary, 234, 235, 237, 239,
 256

perlocutionary, 239, 240

regulative, 285

utterance, 234, 235

Spinoza, Baruch, 32, 42–43, 99

Stalnaker, Robert, 246

Stevenson, Charles L., 286

Strawson, P. F., 109, 112, 246

Stroud, Barry, 181, 277

Subjectivism, 84–85, 95, 114,
 115, 166

T

Tarski, Alfred, *see* Truth

Taylor, Charles, 109, 110, 118–119,
 120

Thomasius, Christian, 99

Transcendental

transcendental argument

forms, 108, 109–111,
 124

in contemporary practical
 philosophy, 181, 182

in contemporary theoretical
 philosophy, 181

in discourse ethics, 264

- and linguistic
 - communication, 233, 253
 - performative or retorsive, 192
 - standard criticisms of, 108
 - transcendental deduction, 10,
 - 18, 31, 85, 91, 108,
 - 112, 114, 131, 133,
 - 136, 153
 - transcendental freedom, 35–54
 - transcendental idealism, 32, 57,
 - 65, 68, 91, 132, 133, 143,
 - 155, 166, 276
 - transcendental methods, 104,
 - 108–109, 111, 114,
 - 124, 127, 129, 134–136,
 - 138, 139, 154, 155,
 - 156, 175
 - analytic (regressive) *vs.*
 - synthetic (progressive), 8,
 - 134, 135, 139
 - transcendental philosophy, 1, 2,
 - 3, 8, 11, 32, 56, 72–82,
 - 83–104, 107–126, 127,
 - 128, 129, 142–152,
 - 159–179, 263
 - critiques of, 163
 - transcendental reflection, 182,
 - 209–218, 224
 - transcendental pragmatics
 - of communication, 209
 - (*see also* Linguistic
 - communication,
 - fundamental conditions of)
 - See also* Apel; Cohen; Fichte;
Hegel; Heidegger; Husserl;
Kant; Løgstrup
 - Truth
 - coherence theory, 266, 267
 - consensus theory, 215,
 - 263–270, 274, 276,
 - 279, 280, 282
 - correspondence theory, 266
 - criteriologically relevant
 - explications, 265, 272, 273,
 - 278, 282
 - evidence theory, 266, 267
 - meaning of truth *vs.* criteria for
 - truth, 270
 - metaphysical-realistic
 - conception, 265,
 - 274, 275
 - regulative idea, 276–279
 - Tarski's logical-semantic
 - explication, 266
 - transcendental-pragmatic
 - conception, 263–282
 - universal-pragmatic
 - conception, 268
 - Tschirnhaus, E. W., 99, 100
- V
- Validity
 - of empirical judgments, 289, 295
 - of moral judgments, 289, 295
 - of moral norms, 289, 290,
 - 292, 297
 - social *vs.* normative, 287, 288
 - Validity claims, 209–231,
 - 268–269, 288
 - discursive redemption of, 269,
 - 271
 - Van der Linden, Harry, 139, 156
 - Vico, Giambattista, 31
 - Vico Principle, 145

W

Weißhuhn, Friedrich August, 36, 47

Wellmer, Albrecht, 264, 265,
270–273, 274, 277, 278,
279, 282, 293

critique of consensus theory
of truth, 264–265, 270–273

Wittgenstein, Ludwig, 22, 109,
110, 210

Wolff, Christian, 23, 100