

TRANSCENDENCE IN POSTMETAPHYSICAL THINKING: HABERMAS'S GOD

MAEVE COOKE
UNIVERSITY COLLEGE DUBLIN

Abstract. Habermas emphasizes the importance for critical thinking of ideas of truth and moral validity that are at once context-transcending and immanent to human practices. In a recent review, Peter Dews queries his distinction between metaphysically construed transcendence and transcendence from within, asking provocatively in what sense Habermas does not believe in God. I answer that his conception of “God” is resolutely postmetaphysical, a God that is constructed by way of human linguistic practices. I then give three reasons for why it should not be embraced by contemporary critical social theory. First, in the domain of practical reason, this conception of transcendence excludes by fiat any “Other” to communicative reason, blocking possibilities for mutual learning. Second, due to the same exclusion, it risks reproducing an undesirable social order. Third, it is inadequate for the purposes of a critical theory of social institutions.

I. INTRODUCTION

In his review of *Postmetaphysical Thinking II*, a recently translated collection of essays by Jürgen Habermas, Peter Dews divides Habermas's *oeuvre* into three phases distinguished by level of confidence in the scope and power of communicative rationality.¹ In Dews' account, the third phase starts at the end of the 1980s with *Postmetaphysical Thinking*, the volume to which the book is a sequel.² It is marked by Habermas's new willingness to concede that the vision of communicative rationality driving his critical theory may lack motivating power. He seems prepared to accept that a critical philosophy needs to provide motivating insights into the core of human existence, and the human impulse to transcend the given, and to question whether his vision of communicative rationality can provide such motivation. In the initial stages of the third phase, Habermas merely hints that this may be a deficiency of his postmetaphysical approach.

Philosophy, even in its postmetaphysical form, will be able neither to replace nor to repress religion as long as religious language is the bearer of a semantic content that is inspiring and even indispensable, for this content eludes (for the time being?) the explanatory force of philosophical language and continues to resist translation into reasoning discourses.³

A similar note of caution is evident in a discussion he conducts with philosophers of religion and theologians around the same time, when he observes that the process of critical appropriation of the essential contents of the major religious traditions is still in train and that its results are hard to foresee.⁴

Nonetheless, his evident awareness of a potential problem helps to explain his later sustained engagement with the relationship between postmetaphysical thinking and religion. In his subsequent writings on religion, politics and philosophy, Habermas describes religion as a reservoir of motivating insights

1 Jürgen Habermas, *Postmetaphysical Thinking II* (Polity Press, 2017). Peter Dews, “Review of: Jürgen Habermas, *Postmetaphysical Thinking II*”, last modified November 10, 2017, <https://ndpr.nd.edu/news/postmetaphysical-thinking-ii/>.

2 Jürgen Habermas, *Postmetaphysical Thinking* (MIT Press, 1992).

3 Habermas, *Postmetaphysical Thinking*, 51.

4 Jürgen Habermas, *Texte und Kontexte* (Suhrkamp Verlag, 1991), 141.

from which postmetaphysical thinking can fruitfully learn.⁵ He calls for a secular mentality that is not secularist, by which he means a mentality that is open to learning from religions, and highlights the importance of translating religious insights into a secular language that would make them accessible to everyone, irrespective of religious belief or lack of it. At the same time, he continues to emphasize the importance for critical thinking of ideas of truth and moral validity that are at once context-transcending and “innerworldly”. He attributes to these ideas a transcending power extending beyond all existing human contexts that can be made sense of only within human practices: it is a transcendent power immanent to the human world. This is his thesis of immanent transcendence, which he also refers to as “innerworldly transcendence” or “transcendence from within”.⁶ As Dews puts it, Habermas’s concept of context-transcending validity “does not rely on a divine transcendence which erupts *into* the here and now”.⁷ However, Dews concludes his review of *Postmetaphysical Thinking II* by querying Habermas’s distinction between metaphysically construed transcendence and transcendence from within. He invites us to consider whether this is, as he puts it, “not a distinction without a difference”.⁸ I read Dews as asking whether there is any significant difference between Habermas’s conception of validity as context-transcending and the conceptions of God as context-transcending that are held by many religious believers. Dews writes: “After all, regardless of the direction we portray the transcending movement as taking, it cannot occur at all without a division — and a gap — between our finite, mortal world and a ‘beyond’ of some kind”. His final, provocative sentence is: “we may well begin to wonder in what sense he [Habermas] does *not* believe in God”.⁹

For the purposes of the present argument, I accept Dews’ invitation to think of Habermas’s commitment to the idea of context-transcending validity as a form of belief in God.¹⁰ My answer to his question is: The distinction *does* make a difference. I argue that in the domain of practical reason there are differences between Habermas’s postmetaphysical “transcendence from within” and metaphysical “transcendence from beyond” that impact significantly on the enterprise of critical social theorizing. I contend, furthermore, that Habermas’s postmetaphysical “God” is not one to which critical social theorists should commit themselves.

I give three reasons for this contention. First, Habermas’s particular version of context-transcending validity curtails the process of socio-cultural learning between postmetaphysical thinkers and religious believers that Habermas now regards as part of the “unfinished project of modernity”. Second, it lacks the radically disclosive quality that some early Frankfurt School critical theorists considered an essential ingredient of truth. Third, it leads Habermas to take an agnostic position with regard to the validity of claims regarding the good life for humans (ethical claims in his terminology), with unwelcome consequences for critique of institutionalized authority — religious authority as well as political and other forms.

Before elaborating on these three troubling consequences of Habermas’s postmetaphysical approach to context-transcending validity, it will be helpful to clarify what Habermas means by “metaphysics”.

Metaphysics for Habermas “is the enterprise of framing a comprehensive view of the world and the place of human beings within it, in which cognitive, normative and evaluative perspectives are fused.”¹¹ By contrast, postmetaphysical thinking insists on a separation between these three perspectives. In the 1980s,

5 Jürgen Habermas, *Between Naturalism and Religion* (MIT Press, 2008).

6 Jürgen Habermas, “Modernity: An Unfinished Project”, in *Habermas and the Unfinished Project of Modernity: Critical Essays on The Philosophical Discourse of Modernity*, ed. Maurizio Passerin d’Entrèves and Seyla Benhabib (MIT Press, 1996), 5,17; Jürgen Habermas, *Truth and Justification* (MIT Press, 2003), 10–11; Habermas, *Postmetaphysical Thinking II*, 82.

7 Dews, “Review of: Jürgen Habermas, *Postmetaphysical Thinking II*”.

8 *Ibid.*

9 *Ibid.*

10 I leave aside the important question of whether Habermas’s commitment to communicative rationality, which is based on an empirically supported “rational reconstruction” of idealizing suppositions built into everyday linguistic practices, is analogous to religious *faith*.

11 Dews, “Review of: Jürgen Habermas, *Postmetaphysical Thinking II*”.

around the time of publication of the *Theory of Communicative Action*,¹² Habermas formulated the separation of perspectives as one between three categorially distinct spheres of validity, each with its own logic of justification: the sphere of science, the sphere of law and morality and the sphere of ethical and aesthetic evaluation. What counts as justification in the sphere of science is a matter for scientists and involves an appeal to truth, construed as universal in scope. What counts as justification in the sphere of law and morality is a matter for legal and moral theorists and involves an appeal to moral-practical rightness, construed as both universal in scope and entailing a principle of universalizability. What counts as justification in the sphere of ethics is determined by the norms relating to the good life operative within a particular form of ethical life (*Sittlichkeit*). What counts as justification in the sphere of aesthetics is determined by expert cultures of art critics in particular cultural contexts. In each case, there is a corresponding mode of argumentation. Habermas's discourse theory of truth and moral-practical rightness, first sketched in the 1970s, is a theory of the logic of argumentative justification in the first two spheres.¹³

In this initial version, Habermas reserved the term “discourse” for forms of argumentation that satisfy certain demanding conditions. In discourses, participants necessarily suppose the approximate satisfaction of idealizing conditions relating to access, conduct and the validity-orientation of argumentation: they necessarily suppose, for example, that no relevant voice may be excluded, that every participant must have equal opportunity to speak and that all participants are concerned with the common search for the single right answer. The conditions are idealizing in the sense that they project an ideal that can, at best, be met only approximately in actual practices of communication. Only discourses concerned with questions of truth (theoretical discourses) and those concerned with moral validity (moral-practical discourses) were considered discourses in the strict sense. Other argumentative forms were characterized as “critique”.¹⁴

In developing his discourse theory Habermas's initial focus was moral validity in a narrow sense.¹⁵ This became known as “discourse ethics”.¹⁶ Despite its misleading name, discourse ethics rests on a categorical distinction between moral validity claims and ethical claims. Habermas aligns himself with Kant's attempt to answer the question of what it means to act rightly in a moral sense, while insisting on significant differences between their two approaches.¹⁷ First, discourse ethics is dialogical: norms are valid if they could be vindicated by an agreement reached among participants in real argumentations (guided by idealizing suppositions); by contrast Kant assumes that individuals can test the validity of their maxims of action “monologically”, in isolation from others. Second, it is a de-transcendentalized version of Kantian ethics. To begin with, it de-transcendentalizes reason. It gives up Kant's dichotomy between an intelligible realm comprising duty and free will and a phenomenal realm comprising inclination, subjective motives and political and social institutions. By contrast, discourse ethics posits a relation of productive tension between the intelligible and the phenomenal — between immanence and transcendence. In addition, its *method* is de-transcendentalized. It replaces Kant's transcendental deduction of the moral principle with a formal-pragmatic argument based on the rational reconstruction of necessary presuppositions of argumentation in general.

Notwithstanding these significant differences, Habermas follows Kant in limiting morality to the class of universally justifiable normative judgments, leaving aside matters of “the good life”. Thus, he demarcates ethics, which deals with questions of the good life for humans, from moral theory, which offers an account of the validity of universal norms and principles.¹⁸ Some critics of discourse ethics in its initial

12 Jürgen Habermas, *The Theory of Communicative Action* (Beacon Press, 1984 & 1987).

13 Jürgen Habermas, “Wahrheitstheorien”, in *Wirklichkeit und Reflexion: Walter Schulz zum 60. Geburtstag*, ed. Helmut Fahrenbach (1973).

14 Habermas, *Theory of Communicative Action*, 1984, 42.

15 Maeve Cooke, “Discourse Ethics”, in *The Routledge Companion to the Frankfurt School*, ed. Peter E. Gordon, Espen Hammer and Axel Honneth (Routledge, 2019).

16 Jürgen Habermas, *Moral Consciousness and Communicative Action* (MIT Press, 1990).

17 *Ibid.*, 203–4.

18 *Ibid.*, 196–97.

formulations expressed concern that it leaves aside many kinds of questions that are morally relevant; furthermore, that it is insensitive to particular needs, aspirations and life-experiences. They pointed out that ethical questions are often experienced as more pressing and more difficult than questions of moral justification in the narrow sense and are at least equally in need of argumentative probing.

His subsequent expansion of the category of discourse helped him to respond to this objection. The expanded discourse theory included ethical discourses, concerned with questions of the good life, and pragmatic discourses, concerned with prudential questions of how to act in specific contexts.¹⁹ Later, it included legal-political discourses, in which ethical, moral and pragmatic questions are interconnected, and discourses of application, which seek to determine how abstract moral principles and norms should be applied in particular cases.²⁰

In the expanded version, ethical validity claims, too, may be the subject of argumentative thematization in discourses. Furthermore, Habermas acknowledges that ethical questions, like moral questions, carry a sense of obligation and may have a context-transcending reference point.²¹ As examples of ethical claims, we could think of claims to the validity of certain fasting prescriptions, or to the validity of particular rules for slaughtering animals, as are common in certain cultures. For Habermas, ethical discourses do not rest on the idealizing supposition that a rational consensus as to the single right answer is achievable. On a pluralist understanding of ideas of the good life, which for Habermas is an integral part of the modern world-view, there is no single right answer to ethical questions; hence, no universal consensus is discursively achievable, even under optimized justificatory conditions. Thus, Habermas's original distinction between "discourse" and "critique", and accompanying distinction between morality and ethics, persists within the expanded category of discourses. On one side, now, there are discourses concerned to thematize pragmatic, ethical or legal-political matters, or to apply laws, ordinances and policies appropriately through reference to context-specific norms. On the other side, there are discourses concerned to justify the truth of propositions, and of de-contextualized moral norms or principles, through reference to an idea of universally binding validity.

The separation of human thinking and action into three distinct spheres of validity, which Habermas likewise considers an integral part of the project of modernity, corresponds to a rejection of the pre-modern, substantive conception of reason. Habermas contrasts substantive rationality, which he associates with religious and metaphysical world views, with the formal understanding of reason that he attributes to Kant and which he sees as gaining traction from Kant onwards.²² When conceived along Kantian lines, reason, at least in the domain of law and morality, is conceived not in terms of its material content but as a framework of formative principles; the focus is on procedure, the conduct of action in line with principles of reason, rather than on what reason concretely tells us to do. Consequently, in embracing a formal-procedural rather than substantive conception of rationality, postmetaphysical thinking abstains from offering substantive ethical orientation and guidance: it does not provide concrete direction with regard to questions of the good life. As Habermas writes, postmetaphysical philosophy gives up its "enlightening role" with regard to life practices as a whole.²³

Habermas rejects metaphysical thinking, not just because it offers a comprehensive view of the world and the place of humans within it, thereby affirming an anachronistic conception of substantive reason; he also rejects it because comprehensive views are underpinned by projections of a transcendent power that is "other" to human reason. In the domain of practical reason, postmetaphysical thinking, at least "for the time being", rejects any notion of validity that has a source beyond the human world of linguistic communication. Note this stronger version of this is Habermas's thesis of immanent transcendence. Distancing himself from the idea of an "Other" to reason, he advocates a deflationary interpretation of

19 Jürgen Habermas, *Justification and Application: Remarks on Discourse Ethics* (MIT Press, 1993), 1–18.

20 Jürgen Habermas, *Between Facts and Norms* (MIT Press, 1996).

21 Habermas, *Justification and Application*, 5.

22 Habermas, *Moral Consciousness and Communicative Action*, 3–4.

23 Habermas, *Postmetaphysical Thinking*, 14–15. Maeve Cooke, "The Limits of Learning: Habermas's Social Theory and Religion", *European Journal of Philosophy* 24, no. 3 (2016).

the “unconditioned” or “absolute”, according to which the transcending power of reason has its origins within the forms of communication through which human beings reach an understanding with one another. He writes:

The linguistic turn permits a deflationary interpretation of the “wholly Other”: [...] In the forms of communication through which we reach an understanding with one another about something in the world and about ourselves, we encounter a transcending power.²⁴

This is the linchpin of his theory of communicative action and the account of rationality corresponding to it. The theory aims to show that a potential for freedom, which is construed in terms of human practices of communication, can be extracted from analysis of everyday language use. In its simplest terms, communicative action is a form of human linguistic interaction that involves raising and responding to validity claims with the aim of reaching mutual understanding.²⁵ Corresponding to the three spheres of validity that constitute rationality in modernity, validity claims may be raised in the sphere of objective knowledge (truth claims), in the sphere of law and morality (moral-practical claims) and in the sphere of evaluative expressions, beliefs and judgments.

Communicative action establishes a relationship between speaker and hearer that is based on a number of normative expectations and obligations. Speakers take on an obligation to support their claims with reasons, if challenged, and hearers take on a similar obligation to provide reasons for their “yes” or “no”. Speakers and hearers seek mutual understanding, in the sense of agreement as to the validity of the claim in question. From this we can see that communicative action is a more or less rudimentary form of argumentation. Corresponding to this Habermas proposes communicative rationality as a conception of context-transcending reason based on the idealizing suppositions built into the very concept of argumentation — suppositions relating to access, to conduct and to the validity orientation of argumentation. Since communicative rationality is based on potentials built into human practices of argumentation, its context-transcending power can be experienced only within such practices and, in the case of moral-practical validity claims, has its source within them. This accounts for the *immanent* character of its transcendence. Its *transcending* power resides in the idealizing suppositions.²⁶ It derives in part from the suppositions relating to access to argumentation and to its conduct, and in part from the supposition, in the case of truth and moral validity, that participants seek to reach agreement on the universally binding character of the proposition, norm or principle under discussion. It should be noted that participants regard the sought-for agreement fallibilistically: they acknowledge that any agreement reached in actual human practices is always open to challenge, even when it is reached under seemingly optimal justificatory conditions.

According to Habermas, this “weak proceduralist understanding of the “Other” preserves the fallibilist as well as the anti-skeptical meaning of the ‘unconditioned’”.²⁷ We may ask, however, whether this is sufficient. In the following sections I focus on some unwelcome consequences.

II. POSTMETAPHYSICAL THINKING AND MUTUAL LEARNING

In this section I argue that Habermas’s immanent-transcendent conception of moral validity impedes mutual learning between the postmetaphysically thinking sons and daughters of modernity and their metaphysically thinking siblings. Diverging somewhat from Habermas’s use of the phrase,²⁸ by “sons and daughters of modernity” I mean those inhabitants of modernity who differentiate between the standards of validity operative in the domain of theoretical reason and those operative in the domain of practi-

24 Jürgen Habermas, *The Future of Human Nature* (Polity Press, 2003), 10.

25 Maeve Cooke, *Language and Reason: A Study of Habermas’s Pragmatics* (MIT Press, 1994).

26 Cooke, *Language and Reason*, 147–66.

27 Habermas, *The Future of Human Nature*, 11.

28 This is a less demanding characterization of the normative horizon of modernity than Habermas offers. See my remarks above on the separation of value spheres and also the demands Habermas makes on modern religious believers in Habermas, *Between Naturalism and Religion*.

cal reason (while allowing for their interpenetration), and who have, in addition, internalized modern normativity with regard to democratic values of liberty, equality and solidarity. For the purposes of the present discussion I follow Habermas in characterizing the postmetaphysically thinking inhabitants of modernity as religious unbelievers and their metaphysically thinking counterparts as religious believers. I acknowledge that this is contentious: not all modern religious believers are metaphysical thinkers in Habermas's sense and not all modern metaphysical thinkers are religious believers.

The problem of mutual learning arises from Habermas's postmetaphysical approach to context-transcending validity. As mentioned, in the domain of practical reason postmetaphysical thinking supports a conception of context-transcending validity only in the case of moral norms and principles. This is because context-transcending validity is tied to the idea of a universal consensus regarding the universalizability of interests: it is tied to the idea that a norm to be valid, must be universalizable, acceptable to everyone, everywhere as being equally in everyone's interests. The postmetaphysical character of moral validity resides in the "innerworldly" constitution of its transcendent quality (as we have seen, it is innerworldly — immanent — because it has its source in human practices). More precisely, the immanent character of moral validity is due to its construction in idealized human practices of argumentation (in this sense it is a constructivist conception). Habermas *defines* moral validity as an agreement reached argumentatively in an idealized communicative situation. The validity of moral norms is not just *tested* in (an idealized) procedure of argumentation, it is *generated* within (an idealized) procedure of argumentation. It does not matter that such a situation is an idealization of actual human practices of argumentation. Indeed, Habermas emphasizes that the "ideal speech situation" is a "methodological fiction", not a condition that could ever actually be achieved.²⁹ What matters is that the very concept of moral validity is defined in terms of this idealizing projection. The "ideal speech situation" is a *conceptual* thought-experiment. For the purposes of conceptualizing moral validity, it calls on us to imagine a social condition in which disputing parties arrive at norms and principles that are morally valid in an unconditioned, universally binding sense.

In Habermas's original formulation of discourse theory in the 1970s, both the concepts of propositional truth and of moral validity were defined in terms of an (idealized) discursively reached agreement.³⁰ From the 1980s onwards, in response to critics, Habermas began to revise his theory of propositional truth. He gradually distanced himself from his previous definition of truth as the outcome of a discursive procedure, replacing it with an idea of truth as justification-transcendent, in the sense that it does not coincide even with the outcome of an idealized justificatory procedure: even in his conceptual thought experiment, in which ideal justificatory conditions actually obtain, an argumentatively reached agreement merely *points towards* truth in an unconditioned, universally binding sense. Habermas writes that it *authorizes* truth. In short, in this new version, there is a gap in principle between truth and justification. Although they remain internally connected (justification under ideal conditions "authorizes" us to refer to something as true), the concept of truth transcends the concept of justification, no matter how idealized. For Habermas, moral validity lacks this justification-transcendent character.³¹ An idealized discursively reached agreement does not merely authorize the rightness of moral norms and principles: it *warrants* their rightness. In Habermas's words: "[i]dealized warranted assertibility is what we *mean* by moral rightness...it exhausts the meaning of normative rightness itself"³² In sum, by contrast with truth, which relates to an objective world deemed to have some essential independence of human practices of justification, the very domain of moral validity is humanly, indeed argumentatively, produced.³³

Habermas remains adamant that moral validity claims are truth-analogous. They have a cognitively construed context-transcending power that derives from their connection with unconditioned, universal validity. In the domain of practical reason only *moral* norms and principles have a cognitive meaning in

29 Habermas, *Between Facts and Norms*, 322–23.

30 Habermas, "Wahrheitstheorien".

31 Habermas, *Truth and Justification*, 237–75.

32 *Ibid.*, 258.

33 *Ibid.*, 262.

this strong sense. As discussed, Habermas adopts an agnostic position with regard to the question of the context-transcending power of ethical validity claims. Certainly, his discourse theory enables criticism of ethical discourses from the point of view of access and conduct. In other words, it enables criticism of ethical validity claims from the point of view of *the way in which* they are thematized, for example, criticism of the exclusion of relevant voices from discussion or of the suppression of some participants' voices. However, it has nothing to say about the validity of their *propositional contents*. The same holds for religious validity claims; in this case, however, by contrast with ethical validity claims, Habermas denies the possibility not only of discursive vindication but also of *thorough-going* discursive examination.³⁴ This has worrying implications for the ability of critical social theory to learn from beliefs, practices and traditions that are justified through appeal to ideas of the good for humans, the source of whose validity is deemed to have some essential independence of human communicative practices.

In Habermas's critical social theory, learning means socio-cultural learning and has a strong cognitive sense. It is a movement in the direction of truth or moral rightness. Moreover, learning means *mutual* learning. Participants in processes of socio-cultural learning engage with their interlocutors as partners in the search for the single right answer to questions in the domains of truth or moral validity. If we probe this conception of learning, we see that, qua mutual learning, it presupposes a shared understanding of the meaning of learning and, hence, a shared conception of truth or moral validity. If participants in argumentation have fundamentally different conceptions of context-transcending validity, and by extension learning, they will not be able to see the outcome of their deliberations as *mutual* learning; at best, they will be able to say that they have learnt something of value for themselves. Think of an exchange between you and me on the question of marriage irrespective of gender. Let's say, my view of moral validity is utilitarian (I might hold, for example, that a moral norm is valid only if it maximizes happiness). Yours is religious (you might hold, for example, that a moral norm is valid only if it is in line with current Roman Catholic teachings). In argumentative exchanges with each other, both of us might change our views with regard to marriage irrespective of gender and, indeed, on the validity of a certain understanding of utilitarianism or of Roman Catholic teaching; the substance of our new views might even converge — for example, we might end up agreeing that marriage irrespective of gender is morally acceptable. However, none of this is sufficient for the result to count as *mutual* learning in the strong cognitive sense in which Habermas understands it. In order for it to count as mutual learning in this strong cognitive sense, you and I, by way of our argumentative exchange, would also have to learn something with respect to the very *concept* of context-transcending validity. The same holds for argumentative exchanges between postmetaphysical thinkers who share Habermas's constructivist understanding of moral validity and metaphysical thinkers for whom context-transcending validity has its source external to human communicative practices. For the parties in the argumentative exchange to regard the outcome as *mutual* learning in Habermas's strong cognitive sense, they would also have to engage with the arguments for his constructivist understanding vis-à-vis a metaphysical understanding, and hold that they had learnt something about the strengths and weaknesses of the respective arguments. In other words, in order for the participants in an argumentative exchange to conceive of the outcome as mutual learning in the strong cognitive sense in which Habermas understands learning, they must also seek a common understanding of what context-transcending validity means. But this implies a readiness on the part of postmetaphysical thinkers to learn from metaphysical thinkers, in this case, from metaphysically thinking religious believers, as regards the validity of postmetaphysical thinking (and vice-versa). Habermas's account of postmetaphysical thinking seems to rule this out *by fiat*. Learning from religion, as he understands it, is a matter of appropriating the propositional contents of religious teachings within a staunchly postmetaphysical framework. He speaks of "critical appropriation" of the contents of religious beliefs,

34 Habermas, *Between Naturalism and Religion*, 129, for a critique see Maeve Cooke, "Violating Neutrality? Religious Validity Claims and Democratic Legitimacy", in *Habermas and Religion*, ed. Craig Calhoun, Eduardo Mendieta and Jonathan VanAntwerpen (Polity Press, 2013).

practices and traditions, of a methodological atheism/ agnosticism with regard to the contents of religious traditions, and of “salvaging” these contents.³⁵

Habermas characterizes learning from religion as a process in which the insights of particular religious traditions are translated into a secular vocabulary that would make them accessible to those with different religious beliefs as well as those with none. Put differently, he views the major world-religions as semantic reservoirs, which secular modern societies may draw on productively in order to enrich their moral vocabularies; however, the religious contents in question must first be translated into a secular language in order to make them accessible to all members of society, irrespective of religious belief. His concern is not just accessibility: the underlying point is that only secular translations of religious utterances are open to thorough-going discursive examination and vindication, since only secular translations have a relation to context-transcending validity in the postmetaphysical sense embraced by Habermas.³⁶ My claim, in sum, is that Habermas advocates learning from religion, but what he has in mind is a circumscribed form of learning that does not extend to the constructivist conception of moral validity affirmed by his postmetaphysical project.

Habermas is committed to the view that only a constructivist understanding of moral normativity is appropriate for the sons and daughters of modernity. As he puts it on occasion, modernity must generate its own normativity.³⁷ If it does not, it will undo the historical learning process set in train within modernity, which has enabled the rational contestation of established authorities and led to a widespread commitment to universalist values of inclusion and equality. But the view that modernity must generate its own normativity *imposes* a certain view of normativity on its inhabitants. It leaves no space for reflective examination and discussion of the question of whether the source of normativity is human or non-human. I see this as a dogmatic closing of the horizons of modernity, out of tune with Habermas’s insistence that modernity is an unfinished project. Indeed, Habermas’s objections to the Hegelian philosophy of history suggest that he sees the project of modernity as not just unfinished but as *unfinishable* project. One of his objections is that Hegelian philosophy of history injects into its reading of history precisely the normativity it seeks to extract from historical processes.³⁸ I read him as objecting not only to the circularity of justification; I take him also to object to how this precludes theoretical re-interpretations of the posited telos of history. Specifically, it precludes theoretical re-interpretation of Hegel’s understanding of the meaning of genuine human freedom, since the truth of the idea of freedom is determined by Hegel’s theory prior to all human action in the world. But, absent the guarantees provided by Hegelian philosophy of history, we cannot assume that the theory’s interpretation of its basic normative concept is the right one, or even that the theory is right to be guided by any version of this concept. It may turn out, for example, that the theory’s interpretation of its basic normative concept, or the very concept itself, serve to maintain and reproduce a kind of social order that is undesirable in light of new ecological visions of how humans should live in relation to themselves and other organisms. By contrast, when the project of modernity is thought of as unfinishable, the meaning and value of freedom and, more generally, the specific contents of what constitutes social learning, must remain a perpetually open question for critical social theory. This would mean that the project of modernity itself is permanently open to re-imagination and re-articulation — allowing even for the possibility that modernity is inherently hostile to the development of genuine human freedom or that human freedom is not the goal for which we should be striving. The same holds for postmetaphysical thinking, which Habermas sees as indispensable for the project of modernity. If critical social theory is to keep open the horizons of modernity, taking seriously the view that modernity is an unfinishable project, it must be open to the possibility of learning about the limitations of postmetaphysical thinking in general and of a constructivist understanding of moral validity in

35 Habermas, *Texte und Kontexte*, 163–139 Habermas, *Postmetaphysical Thinking*, 14–15; Habermas, *Between Naturalism and Religion*, 209–48, for a critique Maeve Cooke, “Salvaging and Secularizing the Semantic Contents of Religion: The Limitations of Habermas’s Postmetaphysical Proposal”, *International Journal for Philosophy of Religion* 60, no. 1–3 (2007).

36 Cooke, “Violating Neutrality?”, Cooke, “The Limits of Learning: Habermas’s Social Theory and Religion”.

37 Habermas, *The Philosophical Discourse of Modernity*, 7.

38 Habermas, *Between Facts and Norms*, 2.

particular. With regard to the latter, postmetaphysical thinking must be ready to learn from its religious (and non-religious) interlocutors whose conceptions of context-transcending validity attribute to it a moment of radical otherness to the ideas of rationality and truth inscribed within any human practices.

III. POSTMETAPHYSICAL THINKING AND TRUTH

A second, related, problem with Habermas's postmetaphysical conception of transcendence is that it limits truth's power to radically disrupt human thinking and behaviour. For the early Frankfurt School critical theorists, commitment to the radically disruptive power of truth marked a crucial distinction between critical theory and pragmatist social philosophy. Acknowledging that both critical theorists and pragmatist social philosophers like Dewey are committed to the endeavour to realize better forms of human life, Horkheimer sees the pragmatists as insufficiently attentive to the ways in which the rationality prevailing within the established social order not alone is hostile to human freedom and happiness but prevents its inhabitants both from seeing this and from imagining what a better form of life would be.³⁹ In other words, for Horkheimer and his Frankfurt School colleagues, critical social theory runs the risk of contributing to the reproduction of an enslaving and degrading social order, if it does not subscribe to a conception of reason, and concomitant idea of truth, that is radically "other" to prevailing conceptions of human rationality and the practices in which they are inscribed. In this respect Horkheimer underscores the importance of a "dangerous, explosive" conception of truth.⁴⁰ We could say: Horkheimer calls for a context-transcending conception of reason that is *radically disclosive* — one that not only transcends the values appealed to in human practices, but has, in addition, the power to open our eyes to ideas of the good life that are decisively different to those currently available.⁴¹ Furthermore, such disclosure is enlightening, exposing the falsity of the ethical practices we engage in in our everyday lives, and of the commitments and convictions structuring and shaping them. Radical disclosure implies an idea of truth or reason that is not purely formal: if denied all content, it would be unable to impact forcefully on us, compelling us to see the falsity of existing practices and forms of life and enabling us to envision different, better ones. While Habermas holds that the transcending power of universal validity claims is a "critical thorn" that sticks in the flesh of social reality,⁴² his discourse theory of practical validity lacks a radically disclosive moment. As discussed, his theory does not say anything about the power of ethical or religious validity claims and its account of the power of moral validity claims construes it purely formally and procedurally as (idealized) discursively achieved universalizability.

Interestingly, in a much earlier article Dews suggests a criticism of Habermas's thinking along these lines.⁴³ Drawing on the writings of the French psychoanalyst Jacques Lacan, Dews hints at certain difficulties with Habermas's particular paradigm of intersubjectivity. He applauds Lacan's alternative intersubjective paradigm for its conception of truth as a power that "transcends the conceptual grasp of finite human subjects", a power that, in consequence, is not theoretically retrievable.⁴⁴ While both Habermas and Lacan conceive of truth in context-transcending, universalist terms, Habermas makes truth the product of communicative reason, at least in the moral-practical domain. By contrast, truth for Lacan has a moment of radical otherness that enables it to present itself to us as a problem.⁴⁵ We could say: for Habermas truth (in the domain of practical reason) is the outcome of a problem-solving discursive

39 Max Horkheimer, *Critical Theory* (Continuum, 1972), 3.

40 Max Horkheimer, "On the Problem of Truth", in *The Essential Frankfurt School Reader*, ed. Eike Gebhardt and Andrew Arato (Urizen Books, 1978), 425.

41 Maeve Cooke, "Contingency and Objectivity in Critical Social Theory: Horkheimer and Habermas", in *Facts and Values: The Ethics and Metaethics of Normativity*, ed. Giancarlo Marchetti and Sarin Marchetti (Routledge, 2016).

42 Habermas, *The Philosophical Discourse of Modernity*, 322.

43 Peter Dews, "The Paradigm Shift in Communication and the Question of Subjectivity: Reflections on Habermas, Lacan and Mead", *Revue Internationale de Philosophie* 49, no. 194 (4) (1995).

44 Dews, "The Paradigm Shift in Communication and the Question of Subjectivity", 490.

45 *Ibid.*, 499.

procedure; by contrast, truth for Lacan is *itself a problem*. This gives it an “imperious” quality.⁴⁶ Lacan writes: “One is never happy making way for a new truth, for it always means making our way into it. It demands that we put ourselves out”⁴⁷. Indeed, for Lacan, the claim of truth is so strong that it can engrave itself in our bodies in the form of symptom.⁴⁸ The “God” that Dews now finds implicit in Habermas’s thinking makes no such claims on us. This postmetaphysical “God” lacks the radically irruptive quality that truth has for Lacan and, apparently, for Horkheimer. Is Horkheimer right to hold that critical social theory needs such a radically disclosive conception of truth if it is to avoid perpetuating social conditions that are hostile to human freedom and happiness? Our discussion in Section I provides grounds for thinking that he is; moreover, for why Habermas should acknowledge this. Appealing to Habermas’s view of modernity as an unfinishable project, I suggested that critical social theories should acknowledge the importance of permanent re-imagination and re-articulation of their basic normative concepts and show willingness to abandon them if they do not contribute towards achieving better forms of human life. Such re-imagination and re-articulation raises two interrelated questions. The first is a question of motivation: What impels those guided by the theory to re-imagine and re-articulate its basic concepts? The second is a question of justification: What allows them to think that their re-imaginings and re-articulations are more conducive to better forms of human life than the conceptions they have superseded? Without an idea of truth as having a content that impacts forcefully on us, making us see the falsity of existing practices and forms of life and helping us to envision alternative, better ones, it would be hard even to begin to answer either question.

IV. POSTMETAPHYSICAL THINKING AND INSTITUTIONALIZED AUTHORITY

A third reason for querying Habermas’s postmetaphysical conception of context-transcending validity arises from its abstinence with regard to the question of the power of ethical validity claims. As we saw in Section I, Habermas maintains that his constructivist approach is appropriate only in the case of highly abstract moral norms, adopting a position of abstinence with respect to the validity of ethical ideas and values. This abstinence has unwelcome implications for critical theorizing about social institutions in general, and religious institutions as a subset of these, providing a further reason to reject Habermas’s “God” — his particular version of immanent transcendence in the moral-practical domain. For, as I now argue, a critical theory of institutionalized authority requires a substantive, context-transcending idea of ethical validity in order to distinguish between authoritarian modes of institutionalized authority and modes that are authoritative but non-authoritarian.

In my account, authority is an *ethically inflected* power. Authority has the power to structure and shape ethical identities: to form humans as concrete beings, in relation to more or less explicit ideas of the good, through ethical prescriptions and recommendations in specific contexts of judgment, decision and action.

Authority is distinct from dominating power. One important difference is authority’s connection with obligation. The power of authority depends on a sense of obligation on the part of those over whom it is exerted; importantly, it is obligation in the form of *self-obligation*. Thus, authority, unlike domination, has a moment of freedom built into it. Acceptance of authority is always in some sense freely granted: at a minimum, there is voluntary recognition and affirmation of the *bearer* of authority.⁴⁹

⁴⁶ Ibid.

⁴⁷ Lacan, quoted in Dews, “The Paradigm Shift in Communication and the Question of Subjectivity”, 499.

⁴⁸ Dews, “The Paradigm Shift in Communication and the Question of Subjectivity”, 499.

⁴⁹ The question of the moment of freedom involved in acceptance of authority is complicated. In Hobbes’ *Leviathan*, for example, humans in the state of nature contract freely to constitute a sovereign power with absolute authority, driven by their interest in the ethical values it fosters (above all, security and commodious living); whether or not they subsequently agree with the content of specific prescriptions is not relevant from the point of view of freedom (with some exceptions, when it is a matter of life or death). For Rousseau in *The Social Contract*, by contrast, humans do not only contract freely to constitute a sovereign

Despite the conceptual connection between authority and freedom, the exercise of authority may undermine freedom, which I conceive of as a form of ethically self-determining agency.⁵⁰ By “ethically self-determining agency” I mean roughly: agency concerned to work out for itself, in interaction with others, what it means to lead a good life, in ways that are not determined by irrational compulsion or by caprice or random choice and decision-making. I characterize “ethical authoritarianism” as a perverted exercise of authority based on a claim to privileged insight into what is good in an ethical sense for those over whom it is exerted. For example, the educational system institutionalized in a particular social order may incorporate liberal-capitalist ideals of successful identity-development, shaping the identities of students according to values such as the competitive acquisition of material goods. If it, and its officers (school principals, teachers, administrators, etc.), impose these ethical values on students, preventing them from questioning their validity, it is ethically authoritarian in my terminology. Or again, religious institutions may incorporate ideas and values in relation to heterosexuality that impede the efforts of some of their members (broadly understood) to work out for themselves what constitutes a good life. If these institutions and their officers (religious leaders, teachers, administrators, etc.) block thematization and critique of their institutionalized ideas and values, they are ethically authoritarian. In both cases, at issue is not the particular ethical prescriptions issued by the institutions in question, but the institutions’ (implicit or explicit) claim that their authority is unquestionable by those over whom it is exerted. I argue, however, that ethical normativity can (and should) be authoritative without being authoritarian. Moreover, since critical social theories are concerned to identify the pernicious effects of social institutions and, by extension, the features of good social institutions, they must be able to make this distinction.⁵¹

I understand social institutions as supra-individual entities that primarily serve the semantic function of shaping and stabilizing social meanings.⁵² Examples include families, religious bodies, parliaments, Churches, trade unions, the World Bank and sports clubs.⁵³ My general claim is that social institutions incorporate ethical values, which are expressed more or less explicitly in the various prescriptions, recommendations and other norms of thought and behaviour issuing from their operations.⁵⁴ The authoritative-ness of social institutions consists in the reasoned acceptance by those subject to them of these ethically inflected norms. By contrast with authoritarian authority, which undermines ethically self-determining agency, authoritative authority contributes to the formation of such agency; it does so by providing ethical orientation in concrete situations that the subjects in question accept or challenge on the basis of rational reflection. In the case of social institutions, its ethical power is manifested in laws, ordinances, rules, policies, prescriptions, recommendations, doctrines and other norms, though it is often tacit rather than explicitly articulated.

Prescriptions and recommendations are authoritative but non-authoritarian when they are affirmed on the basis of rational reflection by particular human subjects in particular life-situations as important aids to orientation in their endeavours to live an ethically good life, and as powerful motivations to live such a life. The ultimate source of their authoritativeness is not a particular institution or its officers, but the truth of the ethical ideas and values to which the institution and its officers more or less tacitly appeal. This presupposes an idea of ethical validity (ethical truth) that transcends the values incorporated in any particular institution. Authority becomes authoritarian when institutions or their officers present themselves as the unquestionable source of ethical validity, or as unquestionable authorities for transmitting particular interpretations of it, permitting no contestation of the ethical ideas and values manifested in the norms they prescribe or recommend. By preventing contestation, they impact negatively on the

power, their subsequent freedom depends on their agreement with the ethical content of the laws to which they are subject (though they may have to be shown that they should agree).

50 Maeve Cooke, “A Pluralist Model of Democracy”, in *What is Pluralism? The Question of Pluralism in Politics*, ed. V. Kaul and I. Salvatore (Routledge, 2019).

51 Cf. Rahel Jaeggi, “Was ist eine (gute) Institution?”, in *Sozialphilosophie und Kritik*, ed. Rainer Forst et al. (Suhrkamp, 2009).

52 Cf. Luc Boltanski, *On Critique* (Polity Press, 2011).

53 Cooke, “A Pluralist Model of Democracy”.

54 Ibid.

freedom — the ethically self-determining agency — of the individuals subject to these norms: these are hindered in their efforts to work out for themselves, in interaction with others, what it means to lead a good life, in ways that are not determined by irrational compulsion or by caprice or random choice and decision-making.

If social institutions are to exercise power that is authoritative but non-authoritarian, they must be open to transformation in response to the ethical challenges they encounter from those subject to their normativity. These challenges may be directed at various aspects of a particular institution's ethically inflected identity: at its operation, its organization and/or its incorporated ideas of the good life. This means, in turn, that social institutions must see themselves, and be seen by those subject to their normativity, as in a permanent process of construction through contestation: they must recognize the inherent instability of their institutional identities. They must acknowledge, furthermore, that the process of construction through contestation is ethically motivated; driven by a concern to shape the institution's identity through incorporation of particular ethical ideas and values. Since contestation is likely to involve plural and possibly conflicting ethical ideas and values, the process of construction will be agonistic rather than harmonious. Nonetheless, the individuals engaged in contestation will consider themselves part of a common project of construction — as co-authors both of a common good that defines the (unstable) identity of the social institution in question and of their own ethically self-determining agency. In sum, for institutions to be non-authoritarian, yet authoritative, they and those subject to their normativity must engage in a perpetual process of mutual ethical identity-constitution, guided by an idea of ethical validity that transcends the particular ethical ideas and values incorporated in particular institutions. Furthermore, the context-transcending power of ethical validity must be understood both as transcendent of human practices and as substantive rather than formal-procedural. Without such an idea of ethical validity, we could not make sense of its radically disclosive power to enlighten us in matters relating to the good life for humans and to point us in the direction of living such a life; consequently, no distinction between the authoritative and the authoritarian would be possible. Habermas's critical social theory's abstinence with regard to the question of the power of ethical validity claims means that it is unable to make this distinction.

Habermas's discourse theory allows no distinction between authoritative and authoritarian ethical validity claims but it does allow for a form of authority that is non-authoritarian. In this theory, as we have seen, the authority of moral norms and principles resides in their universalizability, which is determined in argumentative processes by the human subjects concerned. Since those subject to the authority of moral norms are also their co-authors, moral authority is not authoritarian.

However, this non-authoritarian account of moral validity is of limited help in critically assessing the contents of claims to ethical validity that are tacitly or explicitly raised by, and within, social institutions. From the critical perspective of Habermas's discourse theory, ethical validity claims can be criticized in just two respects: i) irrespective of content, from the point of view of how they are thematized; ii) with regard to their content, if they infringe against moral norms in the narrow sense. However, his theory offers no possibility for assessing their *ethical quality*. But it is primarily ethical quality that is at stake when the authoritativeness of social institutions is challenged. In the case of social institutions, therefore, Habermas's critical social theory hands over the question of authority to anti-authority theorists on the one side, and authoritarian theorists on the other, leaving no room for a third position in which the exercise of authority is authoritative. In short, its ethical abstinence results in a critical perspective that impedes exploration of the features of the authoritative authority of social institutions, religious and otherwise.

V. CONCLUSION

We are now better placed to answer Peter Dews' question of the sense in which Habermas does *not* believe in God. The answer is: he does *not* believe in God in a metaphysical sense; he believes in a "God" that is constructed by way of human linguistic practices. Thus, even if we characterize his conception of moral validity as a conception of God, we must acknowledge it as a resolutely postmetaphysical concep-

tion.⁵⁵ In my discussion, I gave three reasons for why Habermas's "God" is not one that should be embraced by critical social theory. First, in the domain of practical reason, his postmetaphysical conception of context-transcending validity excludes by fiat any "Other" to communicative reason, thereby curtailing the possibility of mutual learning between the postmetaphysically thinking sons and daughters of modernity and their metaphysically thinking siblings. Second, by virtue of the same exclusion by fiat, it runs the risk of reproducing an undesirable social order. Third, it impedes development of an account of social institutions as the locus for the exercise of authoritative, but non-authoritarian authority.⁵⁶

BIBLIOGRAPHY

- Boltanski, Luc. 2011. *On Critique*. Cambridge: Polity Press.
- Cooke, Maeve. 1994. *Language and Reason: A Study of Habermas's Pragmatics*. Cambridge, MA: MIT Press.
- . 2007. "Salvaging and Secularizing the Semantic Contents of Religion: The Limitations of Habermas's Postmetaphysical Proposal". *International Journal for Philosophy of Religion* 60, no. 1–3: 187–207. doi:10.1007/s11153-006-0006-5.
- . 2013. "Violating Neutrality? Religious Validity Claims and Democratic Legitimacy". In *Habermas and Religion*, edited by Craig Calhoun, Eduardo Mendieta and Jonathan VanAntwerpen, 249–76. Cambridge: Polity Press.
- . 2016. "Contingency and Objectivity in Critical Social Theory: Horkheimer and Habermas". In *Facts and Values: The Ethics and Metaethics of Normativity*, edited by Giancarlo Marchetti and Sarin Marchetti, 60–79. London: Routledge.
- . 2016. "The Limits of Learning: Habermas's Social Theory and Religion". *European Journal of Philosophy* 24, no. 3: 694–711.
- . 2019. "A Pluralist Model of Democracy". In *What is Pluralism? The Question of Pluralism in Politics*, edited by V. Kaul and I. Salvatore. New York, NY: Routledge.
- . 2019. "Discourse Ethics". In *The Routledge Companion to the Frankfurt School*, edited by Peter E. Gordon, Espen Hammer and Axel Honneth, 65–81. New York, NY: Routledge.
- Dews, Peter. 1995. "The Paradigm Shift in Communication and the Question of Subjectivity: Reflections on Habermas, Lacan and Mead". *Revue Internationale de Philosophie* 49, no. 194 (4).
- . 2017. "Review of: Jürgen Habermas, Postmetaphysical Thinking II. Translation by Ciaran Cronin". Last modified November 10, 2017. <https://ndpr.nd.edu/news/postmetaphysical-thinking-ii/>.
- Habermas, Jürgen. 1973. "Wahrheitstheorien". In *Wirklichkeit und Reflexion: Walter Schulz zum 60. Geburtstag*, edited by Helmut Fahrenbach, 211–66. Pfullingen.
- . 1987. *The Philosophical Discourse of Modernity*. Cambridge, MA: MIT Press.
- . 1990. *Moral Consciousness and Communicative Action*. Cambridge, MA: MIT Press.
- . 1991. *Texte und Kontexte*. Frankfurt am Main: Suhrkamp Verlag.
- . 1992. *Postmetaphysical Thinking*. Cambridge, MA: MIT Press.
- . 1993. *Justification and Application: Remarks on Discourse Ethics*. Cambridge, MA: MIT Press.
- . 1996. *Between Facts and Norms*. Cambridge, MA: MIT Press.
- . 1996. "Modernity: An Unfinished Project". In *Habermas and the Unfinished Project of Modernity: Critical Essays on The Philosophical Discourse of Modernity*, edited by Maurizio Passerin d'Entrèves and Seyla Benhabib, 38–58. Cambridge, MA: MIT Press.
- . 2003. *The Future of Human Nature*. Cambridge, MA: Polity Press.
- . 2003. *Truth and Justification*. Cambridge, MA.: MIT Press.

55 See note 10 above.

56 I thank the two anonymous reviewers for helpful comments.

—. 2008. *Between Naturalism and Religion*. Cambridge, MA: MIT Press.

—. 2017. *Postmetaphysical Thinking II*. Cambridge: Polity Press.

Horkheimer, Max. 1972. *Critical Theory*. New York, NY: Continuum.

—. 1978. "On the Problem of Truth". In *The Essential Frankfurt School Reader*, edited by Eike Gebhardt and Andrew Arato, 407–43. New York, NY: Urizen Books.

Jaeggi, Rahel. 2009. "Was ist eine (gute) Institution?". In *Sozialphilosophie und Kritik*, edited by Rainer Forst et al., 528–44. Frankfurt am Main: Suhrkamp.



DRAFT