In his book *Truth and Justification* (2003) Jürgen Habermas characterized his theoretical project as a type of Kantian pragmatism. Kantian pragmatism addresses a lacuna in Habermas’ work that has existed since *Knowledge and Human Interests*, namely, the relative neglect of issues in theoretical philosophy. For while Habermas’ formal pragmatics has obvious relevance for theoretical philosophy, its explicit purpose is to formulate a theory of communicative action which itself is meant to ground a critical theory of society and a discourse/theoretic conception of morality, law, and democracy. Kantian pragmatism, in contrast, directly addresses two theoretical questions that arise in light of the linguistic turn: the epistemological question of realism and the ontological question of naturalism. The first asks how we can secure the notion of an objective reality that is the same for all even if our access to the objective world is always mediated by language. Without an answer to this question, the relativism and contextualism endemic to the linguistic paradigm appears unavoidable. The second asks how the normativity that pervades the lifeworld can be reconciled with the fact that sociocultural forms of life evolve naturally. Although these issues are intertwined, for reasons of space we will only address the former question.

The major innovation of *Truth and Justification* is Habermas’ embrace of the thesis that the question of realism can only be addressed by examining both the ‘horizontal’ communicative relations between subjects and the ‘vertical’ relationship that subjects establish with the world through action. In other words, to establish realism we need not only a theory of communication but also a theory of action and learning in which reference to an objective world is ratified *in practice*. The Kantian pragmatism outlined in *Truth and Justification* thus moves beyond the linguistic pragmatism embedded in Habermas’ theory of communicative action and joins forces with
pragmatism in its classical variant. In this way, Habermas distinguishes himself from other contemporary pragmatists like Rorty, Brandom, Apel, and Putnam whose work generally remains at the level of linguistic representation. While these thinkers are concerned with practices, they are mostly concerned with linguistic cognitive practices and not with the practices and actions by which agents cope with problematic situations in the objective world. But without accounting for the latter practices in addition to the former, the problem of realism, and hence the threats of relativism and contextualism, cannot be satisfactorily addressed.

We can arrive at a more detailed understanding of this situation by examining the new theory of truth outlined in *Truth and Justification*. There, Habermas puts forward what he calls a pragmatic theory of truth as a replacement for his long held discourse-theoretic conception of truth, which equates truth with what is rationally acceptable in an ideal speech situation. He does so because the latter theory cannot make sense of the intuition that truth is an unconditional property ‘that cannot be lost’. By failing to do justice to this intuition, however, the discourse-theoretic account cannot explain how truth claims transcend any given context of justification, even an ideal one. On this construal, contextualism is unavoidable. Habermas avoids this result by arguing that the concept of truth as an unconditional property that cannot be lost, rather than originating in our justificatory practices as the discourse-theoretic conception holds, is actually grounded in the practical background certainties that underlie our action within the world. Because truth is ultimately tied to action contexts, contexts in which the existence of the world is ratified in practice, subjects can be sure that their linguistic representations refer to an objective world that is the same for all. This paper has two goals: 1) to chart in detail how Habermas’ new theory of truth overcomes relativism and contextualism, and 2) to argue for the thesis that Habermas’ specific way of meeting this objective is flawed insofar as he resists relativism and contextualism by yoking truth to an overly
narrow concept of objectivity, one not consistent with the larger pragmatic transformation of his thought.

II

The issue of relativism and contextualism is pressing for Habermas because he accepts the post-Kantian ‘movement towards de-transcendentalizing the knowing subject’ (TJ: 175). On this conception, reason is no longer located in a transcendental consciousness located outside of space and time but rather within everyday communicative and lifeworld practices. But in situating reason amongst the rule-governed communicative practices of the life-world it becomes an open question ‘whether the traces of a transcending reason vanish in the sand of historicism and contextualism or whether a reason embodied in historical contexts preserves the power for immanent transcendence’ (TJ: 84). In this section we shall trace how detranscendentalization can be seen as leading to relativism and contextualism, while in the next we will outline Habermas’ strategy to maintain reason’s power for immanent transcendence.

Transcendental philosophy is famously concerned not directly with objects but with the a priori conditions which provide for the possibility of empirically cognizing objects. For Kant, the objectivity of our knowledge is secure because these conditions are universal and necessary. Peirce, the thinker who Habermas most closely follows on this issue, enacts a de-transcendentalization of these conditions by tracing them back to our capacity to use signs. In undertaking a general critique of the Cartesian assumptions that underlie modern epistemology, Peirce formulates a semiotic theory of mind in which a representation is a representation of an object, not simply by being causally provoked by the object, but by being interpreted as a representation of that object by another representation (its interpretant). ‘In place of the dyadic concept of a linguistically represented world, Peirce installs a triadic concept: the linguistic representation of something for a possible interpreter’. By necessarily taking part in the triatic structure of the sign—a sign stands for
an object to an interpreter in some manner—mental representations themselves become semiotic tokens. Here the basic impulse of the linguistic turn is found: to examine the structure of the mental we can circumvent introspection and psychology generally and instead examine linguistic structures. Perice’s schema is unusual, however, because it requires, in addition to a semantic examination of signs and expressions, a pragmatic one. For if expressions always represent states of affairs in some way for some interpreter, then expressions would not be able to establish a semantic relation to objects if they were not also capable of being employed communicatively. While a truth-conditional analysis of expressions might be appropriate in certain theoretical contexts, for Habermas we must embed the capacity to represent objects in the wider capacity to interpret signs, exchange reasons, and engage in symbolic rule-governed practices in general.

In widening our view in this way, we are not only concerned with ‘empirical judgments, but grammatical propositions, objects of geometry, gestures, speech acts, calculations, logically connected propositions, actions, social relations or interactions—in short, basic types of rule governed behavior in general’ (TJ: 11). In examining such behavior our focus shifts from examining reason’s explicit rule-consciousness to explicating the ‘practical knowledge that makes it possible for subjects capable of speech and action to participate in these sorts of practices’ (TJ: 11). This shift is necessary to avoid the ‘regress of rules’—i.e., the regress generated by the need for higher-order rules to interpret the application of rules. Instead, the explicit deployment of a rule is embedded in the ‘intuitive and habitual know-how’ (TJ: 11) that implicitly informs mastery of these rule-governed practices. In other words, the abilities that make up reason must be seen as embedded in the ‘implicit knowledge . . . [that] supports the totality of the web of basic practices and activities of a community that articulates its form of life’ (TJ 11).

This has ramifications for how we should think of Kant’s categories of the understanding. The categories, which for Kant are rules by which a manifold is reduced to unity in a judgment, are,
after the pragmatic turn, explicit and implicit rules of inference that tie together signs and the symbolically articulated web of basic rule-governed practices. This detranscendentalizes the categories by placing their disclosive function not in a reason conceived as standing over-against the world, but rather in the implicit knowledge that supports the inter-subjective moves that are made in historical languages and forms of life. One can immediately discern here the danger of relativism and contextualism, for ‘if transcendental rules no longer signify something rational outside the world, they mutate into expressions of cultural forms of life and have a beginning in time. As a consequence, we may no longer without qualification claim “universality” and “necessity”’ (TJ: 18).

For Habermas, this result is ‘built into the basic concepts of the linguistic paradigm just as skepticism is built into mentalism’ (POC: 355) because if ‘experience is linguistically saturated such that no grasp of reality is possible that is not filtered through language’ (TJ: 30) and the languages that are the condition of possibility for grasping reality are both plural and historical, then the notion that the fundamental rules built into a language are invariant across languages and forms of life becomes unintelligible. If one further claims that that there is no position from which to think and speak that is not already within the horizon of a historically variable language or form of life, then relativism and contextualism seem unavoidable insofar as there is no way for a subject’s claims (moral or theoretical) to ‘transcend from within’ the world that is disclosed by that language or form of life. There would instead be a plurality of internally consistent linguistically disclosed worlds that form the enveloping horizon for any claim to validity.

III

As we know, after detranscendentalization the conditions that make world disclosure possible are now to be found within the world. While this can be seen as leading to relativism and contextualism, on another interpretation it has the more salutary effect of situating ideality within the world,
specifically within the communicative ground of the life-world. Here, reason unleashes its dynamic and transcending power within communication itself. As Habermas puts it, ‘the transcendental tension between the ideal and the real, between the realm of the intelligible and the realm of appearances, enters into social reality of situated interactions and institutions’ (TJ 84).

Habermas’ theory of communicative reason takes its start from this tension. Once one takes the pragmatic turn, it become apparent that reason is not something one has, something substantial, but rather an ability, something that one can do. Specifically, to be rational, speaking and acting subjects must be able, when challenged, to defend and justify their speech acts or actions with reasons. While subjects are not for the most part engaged in justificatory discourse since they generally exist in a taken for granted lifeworld context, they must, to be rational, be able to engage in such discourse with respect to their speech and action. In general, these performances are ‘more rational the better [they] can be defended against criticism’. For a performance to be subject to criticism, however, it must, at least implicitly, make a claim for itself. Speech acts make a variety of claims, to meaningfulness, truth, normative rightness (i.e., being in accord with a legitimate moral rule), and sincerity, while non-linguistic actions implicitly make a claim to being effective at meeting the conditions necessary for a successful intervention into the world. Based upon the different types of norms at play here, Habermas identifies two broad categories of rationality that are in principle not reducible to one another: cognitive-instrumental rationality and communicative rationality.

Action is instrumentally or purposively rational if the reasons implicitly embedded within it can be used to explain a causally effective goal-directed intervention into the world. This type of action is evaluated by a norm of effectiveness. A speech act, on the other hand, is communicatively rational if the propositional knowledge made manifest in it is used to reach understanding with someone about something. The rationality of this type of action—communicative action—is ‘assessed in terms of the capacity of responsible participants in interaction to orient themselves in relation to validity
claims geared to intersubjective recognition. Communicative reason finds its criteria in the argumentative procedures for directly or indirectly redeeming claims to propositional truth, normative rightness, [and] subjective truthfulness'.

Because communicating subjects can, by orienting their arguments to validity claims, sometimes resolve the disagreements that arise because of their ability to freely take ‘yes’ and ‘no’ positions to each other’s speech acts, it is apparent that the communicative form of rationality is based upon ‘the experience of the non-coercively unifying, consensus-promoting force of argumentative speech’. Habermas argues, controversially, that a philosophical reconstruction of this experience demonstrates that speakers counterfactually presuppose validity claims to guide their communicative interactions. As he puts it, if subjects are to reach understanding they ‘must commit themselves to pragmatic presuppositions of a counterfactual sort. That is, they must undertake certain idealizations—for example, ascribe identical meanings to expressions, connect utterances with context-transcending validity claims, and assume that addressees are accountable, that is autonomous and sincere with both themselves and others’ (BFN: 4). One again, we come across the transcendental tension mentioned above, for validity claims have, according to Habermas, a Janus-face:

As claims, they transcend any local context, at the same time, they have to be raised here and now and be de facto recognized if they are going to bear the agreement of interaction of participants that is needed for effective cooperation. The transcendent moment of universal validity bursts every provinciality asunder; the obligatory moment of accepted validity claims renders them carriers of a context-bound everyday practice. Inasmuch as communicative agents reciprocally raise validity claims with their speech acts . . . a moment of unconditionality is built into factual processes of mutual understanding. (PDM: 322)

Habermas thinks that he can resist relativism and contextualism because this social-pragmatic transformation of reason, while bringing reason down to earth, does not completely liquidate the legacy of transcendental philosophy. Instead of investigating the rational structures of a pure consciousness, a post-pragmatic transcendental philosophy will aim to make theoretically explicit ‘the invariant features’ that structure our communicative and symbolic lifeworld accomplishments.
Here we speak, of course, of the formal pragmatic presuppositions of communication. While these features are a priori insofar as their operation makes ‘world-disclosure’ possible, their universality and necessity must be understood in a Wittgensteinian rather than a Kantian sense.

In accordance with formal pragmatics, the rational structure of action oriented toward reaching understanding is reflected in the presupposition that actors must make if they are to engage in this practice at all. The necessity of this “must” has a Wittgensteinian rather than a Kantian sense. That is, it does not have the transcendental senses of universal, necessary, and nouminal conditions of possible experience, but has the grammatical sense of an “inevitability” stemming from the conceptual connection of a system of learned—but for us inescapable—rule governed behavior. After the pragmatist deflation of the Kantian approach, “transcendental analysis” means the search for presumably universal, but only de facto inescapable conditions that must be met for certain fundamental practices or achievements. All practices for which we cannot imagine functional equivalents in our sociocultural form of life are “fundamental” in this sense. One natural language can replace another. But propositionally differentiated language as such (as a “species endowment”) has no imaginable replacement that could fulfill the same function. (TJ: 86)

With this, we now have the full picture of how Kantian pragmatism tries to weave its way through the Scylla of Kantian transcendentalism and the Charybdis of relativism and contextualism. On the one hand, because after detranscendentalization the structures of reason are embodied within our communicative practices—rather than being located in the reason of a worldless subject—Habermas avoids an Idealist rendering of transcendental philosophy. While this makes it possible to claim that these presuppositions are actually effective within the lifeworld it also seems to generate relativism and contextualism by indexing their validity to historically specific languages and rule governed forms of life. But, on the other hand, we now know that there are fundamental or irreplaceable presuppositions of communication whose force transcends any given context or language. Instead of there being a plurality of internally consistent yet incommensurable linguistically disclosed worlds, the idealizing presuppositions of communicative action—which are invariant or fundamental for all forms of life that we can imagine—transcend these worlds from within, giving us common touchstones within our linguistic practices to adjudicate disputes about the factual and normative features of our world.
IV

While Habermas takes it that the discourse-theoretic strategy outlined above is sufficient to refute relativism, in recent work he has found it to be insufficient for dissipating a thoroughgoing contextualism. The contextualism we have in mind does not focus on the relation between languages and forms of life, i.e., whether the claims made in one language have counterfactual force in other languages. Rather it focuses on whether we can, after the linguistic turn, secure the notion of an objective reality to which these languages are referring. In other words, if it really is true, as Habermas claims, that there is no grasp of reality that is not ‘filtered through language’ then even the assurance that there are universally occurring fundamental practices and norms that all languages and forms of life must presuppose does nothing to answer the question of whether these languages are referring to a mind-independent reality at all. Taking the linguistic turn by replacing ‘the transcendental subjectivity of consciousness’ with ‘the detranscendentalized intersubjectivity of the lifeworld’ (TJ: 30) does not solve the problem, it just radicalizes it. Now instead of a mind and its representations confronting the objective world, we have a community and its intersubjective language confronting the objective world. Even if the subject of the problem has changed, the structure of the question remains: how do we get from inside our ‘conceptual scheme’ (either subjective or intersubjective) to the world outside? Without an answer to this question, anti-realism or idealism seems unavoidable.

The problem at issue comes into focus if we consider the concept of truth. The relevance of truth to the problem outlined above should be immediately apparent insofar as truth, at least in its traditional understanding, does not address the relationship between linguistic items but rather the relationship between language and the world. If in taking the linguistic turn, however, we no longer account for truth in terms of correspondence between a proposition and its object and instead see it as something internal to our linguistic, specifically justificatory practices (since after the linguistic
turn the truth of a belief or sentence can only be established by means of another belief or sentence), then the validity of a truth claim cannot depend upon how things stand in the world independently of how they are linguistically given, but upon whether the truth claim is rationally acceptable or ideally justified for a linguistic community. How is this problematic? Well, if ‘in the linguistic paradigm truths are accessible only in the form of rational acceptability, the question now arises of how in that case the truth of a proposition can . . . be isolated from the context in which it is justified’ (POC: 356). In other words, if we can only give a discourse-theoretic concept of truth, one that indexes truth to ideal or rational justification, then we cannot really explain how the truth of a proposition breaks out of the circle of language and transcends the context in which that proposition is uttered. To say that the communicative use of language commits participants to strong idealizations that aim beyond local contexts does not solve this problem because while truth aims beyond a context it never actually gets beyond justification in some context.

To overcome this problem, Habermas thinks, requires the change of perspective that we canvassed in section one: Instead of just examining our communicative use of language within the space of reasons we must also examine our practical relationship to the world established through action. More importantly, we must examine how these two modalities, discourse and action, are fundamentally intertwined. For, as the classical pragmatists saw, it is only by forgetting the intertwined nature of these two modalities that a ‘mirror of nature’ view of knowledge comes about, thereby provoking the problem of the existence of the objective world.

For pragmatists, cognition is a process of intelligent, problem-solving behavior that makes learning processes possible, corrects errors, and defuses objections. Only if it is severed from the context of experiences connected to actions and of discursive justifications does the representational function of language suggest the misleading picture of thought representing objects or states of affairs. The “mirror of nature”—the one-to-one representation of reality—is the wrong model of knowledge because the two-place relation between picture and pictured and the static relation between a proposition and a state of affairs obscures the dynamics of knowledge accumulation through problem solving and justification. (TJ: 26)
The rest of this paper will follow up on the sentiment expressed in this passage by examining how Habermas’ incorporation of a pragmatic perspective into his theory of truth is able to dissipate the contextualist worries that drive his Kantian pragmatism.

V

Before explicating the pragmatic theory of truth we must first examine the inadequacies of Habermas’ discourse-theoretic conception of truth. This older theory, itself a type of pragmatic theory of truth, starts from the familiar thesis that after the linguistic turn language and reality are inextricably intertwined. This fact rules out metaphysical versions of the correspondence theory because agents ‘cannot step outside the circle of language’ to examine whether their propositions correspond to a language-independent reality. To settle a contested truth claim we cannot simply point to language-independent facts, we must instead justify or give reasons for the truth of that claim. Here, ‘the two-place relation of the validity of propositions is extended into the three-place relation of a validity that valid propositions have “for us.” Their truth must be recognizable to an audience’ (TJ: 100).

The discourse-theoretic conception recognizes, however, that truth cannot be identified with what is rationally agreed upon as true by a community or audience. For we know on pragmatic grounds that there is a distinction between truth and rational acceptability. This is brought out by reflecting on the so-called ‘cautionary use’ of the truth predicate: ‘The cautionary use of the truth predicate—no matter how well justified \( p \) may be, it could nonetheless turn out to be false—can be understood as the grammatical expression of a fallibility that we often experience ourselves while arguing and observe in others when looking back at the course of past arguments in history’ (TJ: 38). For Rorty, who introduced the notion of the cautionary use of the truth predicate, the gap between truth and rational acceptability can never be overcome. What this signals for him is that that our
epistemic goal should not be the achievement of truth but an enlargement of the circle of solidarity, i.e., an achievement of rationally motivated agreements in larger and epistemically more informed audiences. For Habermas, on the other hand, the “cautionary” use of the truth predicate shows that, with the truth of propositions, we connect an unconditional claim that points beyond all the evidence available to us’ (POC: 358). What is the nature of this unconditional claim?

Pragmatists, following Peirce, usually account for the unconditional aspect of a truth claim by defining truth counterfactually as the opinion that would be agreed upon by an infinite community at the hypothetical end of inquiry. Here, while we don’t separate truth from our inquiring practices altogether we don’t reduce it to what is rationally acceptable in the here and now. This view follows directly from the fallibilism expressed by the cautionary use of the truth predicate. Fallibilism claims—assuming our procedures of inquiry are in good order—that while we may be assured that most of our beliefs are true, the truth of any particular belief is always conditional. No matter how well justified, a belief may well turn out to false. Fallibilism is different from skepticism, however, because it posits that we exchange or correct our opinions only when we have reason to do so, when there is ‘real’ or ‘living’ doubt as opposed to the ‘paper’ doubt of Descartes. But the only good reason to correct our opinions is to reach an opinion that we would have no reason to exchange, an opinion that withstands all justificatory challenge and receives the assent of everybody. However, because we are aiming at an opinion that withstands all justificatory challenge, the community that examines the opinion cannot just be the current community. Rather, we must imagine the opinion gaining the assent of an infinitely expanding community using better standards of inquiry. The ‘final opinion’ of this infinitely expanding community will be a true opinion and the real will be the object of that final opinion. As Peirce famously put it: ‘The opinion which is fated to be ultimately agreed to by all who investigate, is what we mean by the true, and the object represented in this opinion is the real’.14
With this concept of truth Peirce builds into his system a notion of ‘transcendence from within’: while truth is embedded in a community’s concrete reason-giving practices insofar as our speech acts make claims to truth, truth transcends and regulates these same-self practices insofar as we have reason for disregarding or revising our opinions only because these practices have an inner connection to an idea of an opinion that in the long run will withstand all attempts at correction or refutation. As we know, Habermas follows Peirce in embedding ideality within our discursive practices. However, he detaches the notion of ideal justification from Peirce’s notion that this can only be achieved at an ideal limit. ‘As a regulative idea, the crucial point of orientation toward truth becomes clear only when the formal or processural properties of argumentation, and not its aims, are idealized’ (POC: 365-6). The basic problem with idealizing the aims of inquiry is this: is it not paradoxical to aim at a situation of complete or conclusive knowledge when finitude and fallibility are constitutive of human knowing? Is it plausible to say that the community of inquirers try to instantiate a situation where there is no further need for justification, interpretation, and inquiry, elements that we think are essential for knowledge and human life in general? Is it plausible to say that the goal of reason is to bring reasoning itself to a standstill? Habermas, in a long passage, outlines his alternative view:

Whoever enters into discussion with the serious intention of becoming convinced of something through dialogue with others has to presume performatively that the participants allow their “yes” or “no” to be determined solely by the force of the better argument. However, with this they assume—normally in a counterfactual way—a speech situation that satisfies improbable conditions: openness to the public, inclusiveness, equal right to participation, immunization against external or inherent compulsion, as well as the participants’ orientation toward reaching understanding (that is, the sincere expression of utterances). In these unavoidable presuppositions of argumentation, the intuition is expressed that true propositions are resistant to spatially, socially and temporally unconstrained attempts to refute them. What we hold to be true has to be defended on the basis of good reasons, not merely in a different context but in all possible contexts, that is, at any time and against anybody. (POC: 367)

For much of his career Habermas thought that this was all that could be said about truth: a true belief is one that is the result of a rational consensus forged in an ideal speech situation, i.e., a situation where all the conditions listed in the passage above are met. In recent years Habermas has
come to doubt whether even this conception of truth can work. Habermas canvasses two problems with this view. The first problem is that it cannot do justice to the fact that ‘[w]hereas well-justified assertions can turn out to be false, we understand truth as a property of propositions “that cannot be lost”’ (POC: 358). A proposition makes a claim to being true not because it is the product of a rational consensus forged under ideal conditions, but rather because it is true. As Habermas puts it, ‘a proposition is agreed to by all rational subjects because it is true; it is not true because it could be the content of a consensus attained under ideal conditions’ (TJ: 101). This intuition is not metaphysical, Habermas claims, but is an aspect of the grammar of truth claims. The second problem is that the epistemic conception of truth seems to be subject to an unsolvable dilemma:

No matter how the value of the epistemic conditions is enhanced through idealizations either they satisfy the unconditional character of truth claims by means of requirements that cut off all connection with the practices of justification familiar to us, or else they retain the connection to practices familiar to us by paying the price that rational acceptability does not exclude the possibility of error even under these ideal conditions, that is, does not simulate a property “that cannot be lost.” (POC: 366)

If we take the former road and completely disconnect truth from justification we undermine the basis for endorsing the commonsense notion that our inquiries and practices of justification are essential for getting closer to the truth. Doing justice to this intuition is one of the main selling points for an anti-realist theory of truth. If, on the other hand, we take the latter road and uphold an epistemic conception of truth, we can only have Rorty’s conventionalist position in which the best we can shoot for is not truth but a continuously expanding solidarity. But then what of our intuition that truth is an unconditional property that cannot be lost?

To solve this problem Habermas will argue that while ‘we cannot sever the connection of truth and justification, this epistemically unavoidable connection must not be turned into a conceptually inseparable connection in the form of an epistemic concept of truth’ (TJ: 38). While discursive justification is the only way to settle contested truth claims and is therefore epistemically connected to truth, this does not mean that truth can be identified with what is justified, even in ideal conditions.
Perhaps there is another way of conceptualizing truth, a way that maintains its unconditional character without assimilating this moment to justification under ideal conditions.

VI

To formulate a theory of truth that can uphold its unconditional character without assimilating it to justification under ideal conditions Habermas takes recourse to pragmatism and its theory of action. Before getting to truth, we must therefore briefly review this theory. One of the hallmarks of classical pragmatism is its opposition to a view that privileges the knowing subject vis-à-vis the acting subject. Instead of seeing cognition as an autonomous accomplishment of the intellect, pragmatism argues that cognition is a derivative practice that arises when routine behavioral processes break down in the face of recalcitrant experiences. When one’s background certainties are shaken and doubt arises, agents must undertake inquiry—i.e., theorizing, experimentation, and confirmatory testing—to solve the problem that is confronted. Here, the agent does not just adjust to the situation, but must creatively engage it, thereby bringing something new into the world (a theory, practice, experiment, invention, etc.). It is here that abstract cognition arises as it has generally been conceived by the tradition. If inquiry is successful, the agent acquires new beliefs (rules of behavior) that reestablish undisturbed behavioral patterns. This acquisition is the basis of learning and experiential growth; for it is through a feedback mechanism that the moves made to solve the original problem are integrated into the background knowledge of the acting agent. For the classical pragmatists, this action cycle is not only something that is undertaken at the level of methodologically controlled scientific conduct but is indeed the basic pattern of experience itself.

Although Habermas appropriates this basic model, he modifies it considerably. Instead of speaking of ‘inquiry’ as emergent from a breakdown in pre-theoretical coping, Habermas speaks of ‘discourse’. Habermas makes this substitution because for him the term ‘inquiry’ carries monological
and instrumental implications that are better left behind. In *Knowledge and Human Interests*—the book where he most intensively examines the pragmatic theory of action—Habermas claims that Peirce, in elaborating the pragmatic maxim, turns his back on certain of his own key insights—his early intersubjective theory of language and semiotic theory of mind—to formulate a theory of action where the forms of inference and inquiry are deduced from the structure of instrumental rationality. As Habermas puts it, Peirce embeds the ‘methodological framework of inquiry’ within ‘the behavioral system of instrumental action’.

For Habermas, on the other hand, when background certainties are put into doubt it is not inquiry per se that commences but the intersubjective discursive justification of problematic truth claims. In this case, the realism characteristic of acting subjects ‘is suspended on the reflexive level of discourses, which are relieved of the burdens of action and where only arguments count. Here, our gaze turns away from the objective world, and the disappointments we experience on our direct dealings with it, to focus exclusively on our conflicting interpretations of the world’ (POC: 370).

Habermas here provides a two-level model in which frustrations in our ‘success-controlled dealings’ with the environment lead to dialogical relations within the community of inquirers. On this view, goal-directed action aimed at success and intersubjective discourse aimed at truth have *internal* feedback relations to one another. As Habermas puts it:

> In the spatial dimension, knowledge is the result of working through experiences of frustration by coping intelligently with a risk filled environment. In the social dimension it is the result of justifying one’s ways of solving problems against the objections of other participants in argumentation. And in the temporal dimension, it is the result of learning processes fed by the revision of one’s own mistakes. If knowledge is regarded as the *function* of such a complex structure, it becomes clear how the passive moment of experiencing practical failure or success is intertwined with the *active* moment of projecting, interpreting, and justifying. (TJ: 26)

Despite this disagreement about *what* emerges from disrupted practices, inquiry or discourse, Habermas accepts the pragmatic point that the *purpose* of this emergence is to establish—on the basis of what has been learned through discursive engagement—new beliefs and behavior certainties that can guide future routine action and communication.
When in the course of the process of argumentation, participants attain the conviction that, having taken on board all relevant information and having weighed up all the relevant reasons, they have exhausted the reservoir of potential possible objections to “p,” then all motives for continuing argumentation have been, as it were, used up . . . From the perspective of actors who have temporarily adopted a reflexive attitude in order to restore a partially disturbed background understanding, the de-problematization of the disputed truth claim means that a license is issued for return to the attitude of actors who are involved in dealing with the world more naively. (POC: 369)

VII

The question is: How does incorporating a pragmatic theory of action into his system of thought help Habermas elaborate a ‘pragmatic theory of truth’? Habermas’ pragmatic theory of truth is based upon the thought that truth not only operates at the level of discourse and justification, but also at the level of lifeworld action. As he puts it, truth is a ‘Janus-faced concept’ that mediates between discourse and the action contexts of the lifeworld, between warranted assertibility and the performative certainties that inform action. Truth can mediate between these different modalities because it is in a sense prior to both. Let us explain.

‘Pragmatism’, Habermas claims, ‘makes us aware that everyday practice rules out suspending claims to truth in principle. The network of routine practices relies on more or less implicit beliefs that we take to be true against a broad background of intersubjectively shared or sufficiently overlapping beliefs’ (TJ: 252). In other words, pragmatism makes us aware that even when action breaks down and specific doubts arise they can do so only against the background of an intersubjectively shared network of stable expectations, beliefs, and behavioral certainties that guide routine action and communication. In pragmatic parlance, doubt is always situated doubt. This network of taken for granted certainties, which remain in place ‘underneath’ the attempt to repair the breach caused by a specific doubt, is the origin of the unconditional concept of truth for which Habermas has been searching.

The practices of the lifeworld are supported by a consciousness of certainly that in the course of action leaves no room for doubts about truth. Problem-solving behavior processes frustrations that occur against the background of stable expectations, that is, in the context of a huge body of beliefs that are naively taken to be true. Actors rely on the certainties of action in their practical dealings with an objective world, which they presuppose to be independent and the same for everyone. And these
certainties in turn imply that beliefs that guide actions are taken to be true absolutely. We don’t walk onto any bridge whose stability we doubt. To the realism of everyday practice, there corresponds a concept of unconditional truth, of truth that is not epistemically indexed—though of course this concept is but implicit in practice. (TJ: 39)

Let us remember the immediate problem we have been trying to solve: To do justice to the unconditional character of truth claims one can either disconnect truth from justification altogether, undercutting the intuition that justification helps us to get nearer to the truth, or one can retain the connection, but at the price that a claim justified in ideal conditions can always be wrong, and so never declared unconditional. An examination of the pragmatic role of truth gives Habermas the means to solve this problem: while the connection between truth and justification is epistemically unavoidable at the level of discourse, insofar as the only way to settle truth claims is through justification, agents, when entering discourse, already possess a practical conception of truth that is not indexed to their discursive and justificatory practices at all. It is not indexed to these practices because the ‘unconditional holding-to-be-true’ that is the origin of this conception of truth operates implicitly in our practices and so below the level of reasons. However, this ‘nonepistemic concept of truth, which manifests itself only operatively, that is, unthematically, in action’ projects into discourse by forming an implicit horizon of expectation for our discursive practices. In so doing, it ‘provides a justification-transcendent point of reference for discursively thematized truth claims. It is the goal of justifications to discover a truth that exceeds all justifications’ (TJ: 39–40).

Because the goal of justification is to discover a truth that always exceeds justification, justification can never be final or complete. This is the origin of the sense that our interpretations and justifications in the space of reasons are always fallible. ‘[T]he fallibilist consciousness that we can err even in the case of well-justified beliefs depends on an orientation toward truth whose roots extend into the realism of everyday practices’ (POC: 370). It is this fallibilistic consciousness, this striving for something that cannot be fully instantiated, which provides the motivation for idealizing the conditions of justification. As Habermas says, the
mode of unconditionally holding-to-be-true [that] is reflected on the discursive level in the connotations of truth claims that point beyond the given context of justification . . . requires the supposition of ideal justificatory conditions—with a resulting decentering of the justification community. For this reason, the process of justification can be guided by a notion of truth that transcends justification although it is always already operatively effective in the realm of action. (POC: 372)

We claimed above that truth is a Janus-faced concept that mediates between lifeworld and discourse, between performative certainty and justified belief. While, as we just saw, it facilitates the movement ‘upward’ of practical certainties into discourse, providing discourse with a justification-transcendent point of reference and a motivation to idealize the conditions of justification, it also ‘allows for a retransformation of rationally acceptable assertion into performative certainties’ (TJ 253). In other words, because truth is operative within action contexts as well as discourse, when argument comes to an end and truth claims are de-problematised, we are licensed to return to an action framework where justified beliefs are transformed into behavioral certainties. In making this transition we do not leave truth behind, but, as it were, return to its native realm. But subjects do not return from discourse to action unchanged. The mode of unconditionally-holding-to-be-true not only informs the realm of reasons insofar as it projects into the discursive level of justification, but it is also informed by this realm since the settlement of problematized truth claims in discourse feeds back into our background practical certainties. Here, the consciousness of fallibilism generated at the level of discourse ‘reacts back upon everyday practices without thereby destroying the dogmatism of the lifeworld’ (POC: 371).

This fact allows Habermas to avoid a criticism that might be directed at his theory of truth: namely, that truth in its practical form is located in a realm that is prior to reason. By this criticism’s lights Habermas’ pragmatic theory of truth is akin to Heidegger’s notion of truth as alethea, which in disclosing a world in advance, is not subject to rational revision. But this criticism misses the mark in two ways. First, Habermas’ theory of truth is formal insofar as the concept of truth refers to the doxic modality of our background beliefs, i.e., our implicitly ‘holding-them-to-be-true-unconditionally’. Here, truth is not a mode of being uncovered, as it is for Heidegger, but ‘a memory’ of
unconditional certainty that structures our action and discourse. Second, because of the downward influence of discourse on action, the content of the practical certainties that underlie action and routine communication can be modified and revised—although, as Sellars once put it, not all at once. In other words, even though the doxic modality of truth as unconditional is maintained the content of the background certainties that make up the lifeworld can be rationally revised in light of new learning.\textsuperscript{18} Now there is available to Habermas the notion of an open future that can be creatively integrated into a form of life without that integration being an arational ‘gift of Being’.

**VIII**

According to the view propounded above, the origin of our concept of truth is to be found neither in the structure of the world itself, nor the rational structure of goal-directed action, nor in the rational dynamics of communicative action aimed at understanding. Rather it is to be found in the practical certainties that underwrite both action and communication, or more specifically, in the certainties that underwrite the pragmatic action cycle, a cycle that includes both of these moments. Even though this conception of truth seems radically anthropomorphic, tracing truth to the modality of our unconditionally holding-something-to-be-true, Habermas thinks that this conception of truth can support realist intuitions. The question is how.

Let us come back to the contextualist doubt that motivated Habermas to take on board the pragmatic theory of action in the first place. This doubt, if we remember, centered on the issue of whether the notion of an independent objective reality that is the same for all can be secured even if access to that reality is always mediated by language. A more specific way of putting this doubt concerns the question of truth, of how the truth of a proposition can transcend a specific context of use, when—after the linguistic turn—truth is conceptually inseparable from justification. We saw above how the pragmatic theory of truth deals with this latter more specific doubt. Truth-claims
within discourse transcend immediate and even ideal communicative contexts because they reflect a 
mode of unconditionally holding-to-be-true that is always already operative in the realm of action. 
But how does this move address the larger issue of securing the notion of an independent objective 
reality? It might seem as if this new notion of truth binds it to a context just as readily as the 
discourse-theoretic theory of truth. In the latter theory, truth, although formally transcendent insofar 
as it is a pragmatic presupposition of discourse, is still tied to a context of (ideal) justification. In the 
newer theory, truth is also formally transcendent insofar as the mode of unconditionally holding-to- 
be-true projects into discourse giving us a notion of truth that is beyond justification. However, 
discourse is still embedded in a context, not of justification, but of implicit beliefs unconditionally 
held to be true. It is unclear how this context is not as structuring of the truth as the ideal 
justificatory context that pertains to the discourse-theoretic concept of truth.

Habermas addresses this problem by intertwining the concept of truth with the concept of 
objectivity. Contextualism is ultimately otiose, Habermas thinks, because the certainties that are the 
origin of our concept of truth are the certainties that acting agents rely on in their practical dealings 
with an objective world, a world ‘which they presuppose to be independent and the same for 
everyone’. In other words, the certainties that are the formal origin of our concept of truth are 
always about the objective world. As Habermas puts the point, ‘[b]ecause acting subjects have to 
cope with “the” world, they cannot avoid being realists in the context of their lifeworld’ (POC: 370).

But why do acting agents presuppose there to be an independent objective world in the first 
place? What is the origin of the concept of objectivity? The concept of objectivity arises ‘through 
failure’, when action breaks down ‘we experience in practice that the world revokes its readiness to 
cooperate, and this refusal gives rise to the concept of objectivity’ (TJ: 255). This experience of failure, 
of our individual and cooperative projects undertaken with others, generates an implicit concept of 
an independent world that is the same for all, which feeds back into our system of background
beliefs, informing our future action and behavior. From now on, within action we ‘presuppose the objective world as a system of possible referents—as a totality of objects not of facts’ (TJ: 27). Like the concept of truth, the concept of objectivity circulates upward into discourse, giving it the assurance that our statements refer to objects that exist independently of our representations. In this way, objectivity becomes a practical presupposition not only of action but of communication as well.

‘As subjects capable of speech and action, language users must be able to “refer” “to something” in the objective world from within the horizon of their shared lifeworld if they are to reach an understanding “about something” in communicating with one another’ (TJ: 89). Because of the intertwined nature of truth and objectivity, there is now an internal connection between our assertions, which make claims to truth, and the world of objects to which these assertions refer.

Because of the internal connection between truth and objectivity at both the level of action and discourse we can be sure that our justification-transcendent truth-claims refer to a mind-independent world. This pulls the rug out from under a contextualism that indexes truth to a particular context, for now truth claims ultimately refer to a single context—i.e., the objective world as a system of possible referents—whose independence is ratified not at the level of linguistic representation but in practice.

IX

While Habermas’ account of truth resists contextualism it does so, in our view, at the price of yoking truth to an overly narrow concept of objectivity. For Habermas, the concept of objectivity that is implicit in our practices only pertains to our interactions with ‘independently occurring processes’
(TJ: 256), i.e., with processes that occur in the objective world. Since truth and objectivity are intertwined, this means that the background beliefs and certainties that are the formal origin of our implicit concept of truth can only be about such processes. But to say this requires denying the holism of the lifeworld certainties that underlies the pragmatic action cycle: While background beliefs that support action and communication about the objective world can have the doxic modality of certainty, other sorts of background beliefs cannot. It is clear that the lifeworld can be retrospectively stratified in a variety of ways. But nowhere does Habermas provide an argument for stratifying it in accordance with the doxic modality of our most deep-seated background beliefs. Unless one provides an argument to the contrary, it is unclear why the dogmatism of the lifeworld only pertains to certainties about the objective world and not to deep-seated assumptions about how to relate to other people, about which emotional responses are appropriate, about what is good or valuable, etc. While at the level of practice Habermas’ concept of objectivity refers to a mind-independent world, the concept of truth reflexively refers to the doxic status of our background beliefs, whatever their content.

What would happen if we pried truth from the specific concept of objectivity that Habermas credits to theoretical praxis? We would still be able to account for the justification-transcendent nature of truth-claims insofar as the projection of our ‘unconditionally-holding-something-to-be-true’ into discourse would give us the notion within our discursive practices that there is a difference between something being true because it is true and our finding something-to-be-true because of its rational acceptance in a discourse. But on this conception our finding something-to-be-true because it is true does not assume that what is true only concerns the facts circumscribed by Habermas’ narrow theoretical construal of objectivity. Separating truth and objectivity in this way would, of course, tempt one to violate Habermas’ Kantian injunction against assimilating theoretical and
practical reason, truth and moral validity. But in our view this would be a virtue of the revised theory rather than a vice. An argument for this, however, will have to wait for another occasion.

NOTES

1 I would like to thank the National Endowment for the Humanities for support in the writing of this paper.

2 Jürgen Habermas, *Truth and Justification*, trans. B. Fultner (Cambridge, Mass: MIT Press, 2003). Hereafter this text will be referred to as TJ.

3 Here Habermas is reviving, in a post-linguistic context, certain insights that he expressed in his early work. See The Positivist Dispute in German Sociology, ed. T. Adorno (New York: Harper and Row, 1976) and Knowledge and Human Interests (Boston: Beacon Press, 1971).


6 This can be done in the case of action because actions—if they are not merely rote—implicitly contain a type of know-how that can be formulated as reasons which rationalize the action.


8 Cognitive-instrumental rationality is itself a species of the larger genus of teleological rationality, which includes, besides instrumental rationality, strategic rationality. While both of these forms of rationality are criticized by the norm of effectiveness, strategic rationality does not pertain to non-linguistic action but to speech acts by which we try to influence the behavior of other actors.

9 Jürgen Habermas, *The Philosophical Discourse of Modernity* (Cambridge Mass: MIT Press, 1992), p. 314. Hereafter this text will be referred to as PDM.

In *Truth and Justification* Habermas couches the issue as one concerning the necessity of the world that ‘appears to be objective to all sociocultural forms of life’. If this necessity cannot be secured then our access to the world might just be an anthropological generalization ‘from species-specific experiences’. To deal with this problem, Habermas is forced to take up the ontological question of naturalism that we in this paper are leaving aside. However, in his article on Rorty (included in POC), Habermas addresses the contextualism issue as a problem endemic to the linguistic paradigm. Here, there is no mention of the world’s necessity or the ontological problem of naturalism. Our discussion will follow this presentation.

Of course, many semantic conceptions of truth (deflationary theories) account for truth in intra-linguistic terms different than those given in the discourse-theoretic account of truth outlined here. Habemas takes it that these theories are flawed because they do not account for the normative force implicit in the disquotational schema. See POC: 361-4.


Another problem with idealizing the aims of inquiry that Habermas does not mention is that of specifying what is meant by the end of inquiry without that specification simply glorifying the standards and canons of reason we currently employ. The situation of fallibilism does not just apply to the truth of specific beliefs but also to conditions of idealization.

*Knowledge and Human Interests*, (Boston: Beacon Press, 1971), p. 134. This, I think, is an overly narrow interpretation of Peirce that relies upon reading Peirce’s 1877-8 papers in abstraction from his later work and self-criticism. Perice, in my view, wants precisely the position that Habermas claims to want, i.e., one that sees our knowledge-constitutive interest as not only including our success-controlled relations to the environment but also our dialogical relations with the community of inquirers. I assume that Habermas’ current reading of Peirce on this issue is more charitable then when he wrote *Knowledge and Human Interests*, a work which still contained many of the prejudices towards pragmatism evinced by the first generation critical theorists.

By claiming that breakdowns in action only lead to discourses aiming at truth, Habermas significantly restricts the meaning of the pragmatic theory of action. It does so in two ways. 1) For pragmatism the feedback loop between action and communication takes place across the spectrum of validity spheres, including the moral sphere. Problems that arise in our coping do not just concern the objective world, but also the moral, aesthetic, and political worlds. Unlike
Habermas, pragmatism does not think that the pragmatic action cycle is primarily for the sake of knowledge. Knowledge is one amongst many ends—and a specialized one at that. 2) Pragmatism has a notion of non-teleological action in which means and ends, parts and whole, are internally integrated leading to ‘an experience’. Success here has a different meaning than when one is aiming at a pre-set end. For more, see Hans Joas, _The Creativity of Action_ (Chicago: Chicago University Press, 1996).

18 For more on this issue, see Nicolas Kompridis, _Critique and Disclosure: Critical Theory Between Past and Future_ (Cambridge, Mass: MIT Press, 2006).

19 The resistance to our cooperative projects with others gives rise to our assumption that the world we are dealing with is the same world our partner is dealing with. This notion, that the world is the same for all subjects, is an essential aspect of objectivity.

20 Besides his pragmatic theory of objectivity Habermas has a complex theory of objectivity that only uses the resources of his theory of communication. For this theory see Christina Lafont, “Is Objectivity Perspectival? Reflexions on Brandom’s and Habermas’s Pragmatist Conceptions of Objectivity,” in _Habermas and Pragmatism_, eds. M. Aboulafia, M. Bookman, and C. Kemp (London: Routledge, 2002).

21 To fully cash out this theory would require a much more detailed account of Habermas’ theory of reference. This would involve describing the ‘referential system’ that is built into a natural language insofar as it contains indexical and deictic expressions. In using these expressions we commit ourselves to the notion that the object that we speak about and act upon remain the same through the various descriptions and predicates that we can ascribe to it. Indeed, without this presupposition, learning would be impossible insofar as our views can change about an object only if we can identify it independently of our descriptions.

22 For some of the ways that Habermas stratifies the lifeworld see “Actions, Speech Acts, Linguistically Mediated Interactions, and the Lifeworld” in POC. The structure of the lifeworld in Habermas is, of course, a complex affair. There are two issues that must be kept separate. First there is the well known tension in Habermas’ account between his stress on the holism of our background knowledge and his urge to claim that in modernity the breaking up of reason into different spheres— theoretical, practical, and aesthetic—reacts back and structures the lifeworld itself. In modernity, do the components of the lifeworld remain ‘fused together’ and ‘split up into different categories of knowledge only under the pressure of problematizing experience’ (POC 244), or are these categories split up this way only because they already exist proleptically in the lifeworld? The second issue—the one that is relevant for us—does not concern the
content of the background beliefs that make up the spheres of reason but their respective doxic statuses. Are our background beliefs split up into the different spheres of reason in accordance with their different doxic statuses—such that only theoretical beliefs have the doxic status which is the origin of truth—or do the different doxic statuses that our beliefs can possess cut across all the spheres of reason? In other words, can beliefs in any sphere of reason have the doxic status of unconditional certainty? Regarding what Habermas calls our ‘deep-seated background knowledge’ (POC: 242) I can find no argument for the view that doxic statuses are botanized in accord with their sphere of reason.

Habermas seems to assume that his view concerning the categorical distinction between labor and interaction, the objective world and the lifeworld, facts and values, etc. are enough to determine that truth can only concern what he counts as objective. But his adoption of the pragmatic theory of action, which *prima facie* puts these distinctions into question, requires an explicit argument for breaking up the lifeworld in accordance with doxic status.