<table>
<thead>
<tr>
<th>Title</th>
<th>The problem of limit concepts in Habermas: toward a cognitive approach to the cultural embodiment of reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Strydom, Piet</td>
</tr>
<tr>
<td>Publication date</td>
<td>2018</td>
</tr>
<tr>
<td>Type of publication</td>
<td>Article (peer-reviewed)</td>
</tr>
<tr>
<td>Link to publisher's version</td>
<td><a href="https://www.pdcnet.org/philinquiry/content/philinquiry_2018_0042_0001_0168_0189">https://www.pdcnet.org/philinquiry/content/philinquiry_2018_0042_0001_0168_0189</a></td>
</tr>
<tr>
<td></td>
<td><a href="http://dx.doi.org/10.5840/philinquiry2018421/29">http://dx.doi.org/10.5840/philinquiry2018421/29</a></td>
</tr>
<tr>
<td></td>
<td>Access to the full text of the published version may require a subscription.</td>
</tr>
<tr>
<td>Rights</td>
<td>© 2018, Philosophy Documentation Center. All rights reserved.</td>
</tr>
<tr>
<td>Item downloaded from</td>
<td><a href="http://hdl.handle.net/10468/7533">http://hdl.handle.net/10468/7533</a></td>
</tr>
</tbody>
</table>

Downloaded on 2019-08-03T15:37:09Z
The Problem of Limit Concepts in Habermas:
Toward a Cognitive Approach to the Cultural Embodiment of Reason

Piet Strydom
University College Cork, Ireland

Abstract
This essay deals with Habermas’ concept of truth in his late theoretical philosophy. Assuming his suggestive yet highly inspiring inauguration of a cognitive turn in Critical Theory, it probes his use of the notion of limit concept against the background of the tradition from which it originally derives with the intention of identifying the notion’s potential for taking this promising departure further. It brings to the fore a number of issues in his late writings that reveal the presence of what may be considered the problem of limit concepts in his thought. For present purposes, these issues are located in two areas: Habermas’ revision of his long-held concept of truth and the related criticism of Peirce; and his account of the role of limit concepts like truth and warranted assertibility or rational acceptability in processes of discursive justification. The analysis finds that there is a structural deficit in his presentation that could be filled by cognitively conceived cultural structures that not only correspond to the major types of limit concepts, but also answer to his undeveloped vision of the ‘cultural embodiment of reason’.

Key words: Aristotle, cognitive turn, Critical Theory, culture, Habermas, justification, Kant, limit concept, Peirce, truth

Introduction
This essay focuses on the decisive role that the notion of limit concept plays in Habermas’ thought, doing so from the perspective of how an investigation of it could help take further his inspiring inauguration of a cognitive turn in Critical Theory. It opens with a preliminary review of his conception and use of the notion of limit concept in both his earlier and later work. The notion is then more thoroughly explored by being placed in the context of the mathematical-philosophical tradition from which it originally derives and investigated by reference to its subsequent adoption by modern philosophers and mathematicians, including Kant and Peirce who are of critical importance to Habermas. Against this background, a critical evaluation is finally offered of a number of issues in his late theoretical philosophy that shed light on what may be considered the problem of limit concepts in Habermas. Attention is given to two selected issues, the first being the hierarchy of truth concepts resulting from the revision of his concept of truth and his related criticism of Peirce; and the second his account of the paradox of justificatory discourse and, following from it, the fallibilist proviso to which all truth claims are subject. The suggestions regarding a cognitive interpretation accompanying this argumentation seek to fill a structural deficit in Habermas’ presentation and thus to answer to his vision of the ‘cultural embodiment of reason’ by translating the notion of limit concept into different types of cultural structures serving as idealized orientation complexes at distinct levels. The structures thus identified are not only presupposed by Critical Theory, but also have the potential of bringing its philosophical and social-scientific sides closer together.

I Limit concepts in Habermas
In his writings, Habermas from time to time explicitly calls on the notion of a ‘limit concept’ (Grenzbegriff), but on the majority of occasions he operates implicitly yet most significantly with this notion. This is the case in particular with his formal-pragmatics and, by extension, with his suggestions regarding what he calls the ‘cultural embodiment of reason’.

1 In English-language philosophy, the alternative expression ‘boundary concept’ is also used – see, for example, Sainsbury (1996) and Varzi (2015). However, the issues discussed here do not appear on their radar screen.
In respect of his explicit use of the notion, first, examples can be drawn from both his earlier and later writings. In a late instance, he writes as follows regarding the concept of world:

‘The world consists in everything we can refer to in true propositions. There’s not much more to say than that. In speaking of the totality of objects of possible true propositions, we refer to all those constraints that challenge us to learn something about the world; whereas in speaking of the totality of facts, we anticipate what we might have learned at the end of time from those challenges about the world. What for us gains the quality of something worth knowing depends not merely on the world but just as much on our position in the world. That is why “the world” is a limit concept or an “idea” (in the Kantian sense): the world consists in all those objects of possible references, not in the language-dependent facts that we come to assert regarding these references’ (2007: 39-40).

Whereas in this quotation Habermas is concerned with the world of all possible objects of experience and reference as the limit concept of the theory of knowledge of the objective domain, in much earlier work done in preparation for his major book on communicative action he deals also with limit concepts pertaining to the moral or normative domain. Since the ethical theory he defends is of a formalistic type that attends only to questions of justice to the exclusion of questions of the good life, he submits:

‘[I]n a consistently thought-through discourse ethics compassion and solidarity appear at least as limit concepts...From this perspective, “compassion”, compassion with past generations’ suffering from the violation of their moral and bodily integrity, is a limit concept of discourse ethics in a similar sense as “nature-in-itself” is a limit concept of transcendental-pragmatic epistemology' (1984a: 515, 517, my translation).

In view of the fact that a completely just human form of life is an inherently contradictory notion, given that its genesis involved the irrevocable suffering of the killed, maimed and tortured and that its universalism is in principle unrealizable, he agrees with Walter Benjamin that this contradiction can be compensated only by extending its universalism by ‘anamnestic solidarity’ (517) – compassionate solidarity with the irremediable suffering of past generations by thinking ourselves into their despair and keeping it alive in our memory.

The sense in which Habermas uses the notion of limit concept in these two cases amounts to marking the extremity or end which, like a wall, the knowledge production process regarding the objective world and the discursive articulation of justice ultimately run up against. Beyond those bounds these two processes do not and cannot go.2 Whereas the process of knowledge production cannot say anything about the world as a totality of objects since it is confined exclusively to establishing the totality of facts about the world, discourse ethics’ confinement to justice or, more broadly, to moral concepts occludes any possibility of it getting embroiled in compassionate relations. Although these qualifications apply, it is nevertheless the case that Habermas considers the limit concept of the world as having the status of ‘an “idea” (in the Kantian sense)’ and that the limit concept of ‘compassion’ is a compelling ‘postulate’, none of which can be ignored. It is indeed so that an idea invokes something of which we have no knowledge and that a postulate is no more than an assumption that something is the case, yet as limit concepts each of them does have a range of definite and indispensable enabling functions.3 Without them neither knowledge production nor discursive justification would be possible or, at best, would unfold in a hopelessly haphazard, fragmentary and highly deficient way.

---

2 It could be argued that there is also a third process, the process of self-cultivation and subject-formation with its own limit concepts, which is attended to below.

3 Legg (2014: 207-10) offers an illuminating list of functions or uses of limit concepts with reference to Peirce.
This leads to Habermas’s less explicit yet nevertheless most significant use of the notion of limit concept. In his account of formal-pragmatics where he operates unthematically with the notion, a probing eye is able to reveal the contours of both the status and role of limit concepts. To begin with, two aspects of his use of the notion are identified here to set the scene for the investigation presented in the following sections of the essay.

First, the concept of world has a defining significance for Habermas’ formal-pragmatics. In its construction, the concept of world becomes differentiated and amplified on different levels into what he considers the three formal-pragmatic domains which are delimited by the ‘formal world-concepts’ of the ‘objective world’, the ‘social world’ and the ‘subjective world’ (Habermas 1984b: 50-1). In turn, each of these three worlds internally accommodates its own particular set of specific concepts – for example, truth and efficiency in the case of the objective world, rightness and justice in the case of the social world, and truthfulness and appropriateness in the case of the subjective world. Second, the three world-concepts are regarded as ‘formal presuppositions’ that are ‘necessary’ in the sense that they are enabling conditions. As such, they are ‘idealizations that orient subjects’ (Habermas 2003: 93) and allow them to refer to something, relate to others and acknowledge each other.

It could be surmised, further, that each of the specific concepts accommodated by the three worlds, such as the concepts of truth, rightness and truthfulness, are also limit concepts and that they are likewise necessary idealized formal presuppositions. That this interpretation is plausible is confirmed by his characterization of them as ‘unconditionality’ (Habermas 2003: 99). These concepts are not simply enabling like the higher-level world-concepts, however, since they more specifically allow engagement with what is being isolated, referred and related to in the different worlds from the corresponding basic orientational attitudes. Simultaneously, they are also vital in facilitating the directing and guidance or recursive regulation of the various practices carrying such engagements. Being unconditional, they do this by taking an abstract form devoid of all specific content. This entails that for their structuring force to be brought to bear on actual engagements and practices, the voicing of what Habermas calls ‘[v]alidity claims...based on formal world-concepts’ (1984a: 50; 1979: 58, 68) is required. Validity claims, in his view (1987: 231, 2003: 99), form part of the lifeworld as mechanisms that mediate between unconditional formal presuppositions encapsulated by culture and actors in spatiotemporally identifiable contexts of social practices. In his late work where he traces the transformation or ‘detranscendentalization’ of Kant’s transcendental philosophy, he adopts a broader perspective on this same set of complex relations when he suggestively speaks of the ‘cultural embodiment of reason’ (2003: 218). In effect, this conception captures the process whereby validity claims activate concepts like truth, rightness and truthfulness so that their rational – or cognitive! – significance penetrates down from their elevated position in the domains marked by the world-concepts to insert their structuring force in linguistically mediated social practices and forms of life.

The conclusion that follows from this preliminary overview of Habermas’ employment of the notion of limit concept is that it is in one respect explicit and quite clear, but in another implicit and thus both fruitfully suggestive and calling for exploration and clarification. Besides arriving at a grasp of how limit concepts figure in his thought, the question raised by the sense that further investigation

---

4 See Habermas (1979: 68) for his original statement.
5 Already here one gets the impression that something structural is lacking between unconditional formal presuppositions and validity claims. Concepts like truth, rightness and truthfulness are not validity claims, as he often intimates, but formal and abstract principles which, while making regulation of the discursive process triggered by a validity claim possible, they do not themselves exercise regulation. Such a recursive function requires an intervening structure distinct from both formal presuppositions and validity claims.
is required is whether his assumptions regarding the status and role of limit concepts and, consequently, his use of the notion entail any identifiable problems. The route followed in the next section to obtain the desired clarification is to undertake a brief foray into the background of the notion of limit concept. This path leads through the venerable mathematical-philosophical tradition.

II The mathematical-philosophical background

The mathematical-philosophical tradition was founded, nurtured and sustained by serious research and design work done by mathematicians, scientists, philosophers and logicians over the past more than 2000 years and has had a thorough-going yet largely forgotten impact on philosophical and social thought. From a core line of analysis running from Aristotle, via Hindu-Arab thought, Fibonacci, Galileo, Wallis, Newton, Leibniz, Berkeley and Kant, to nineteenth-century philosophers and mathematicians such as Cauchy, Dedekind, Cantor and Peirce, it is apparent that the closely related issues of infinity, infinite processes and limit concepts were key problems of this intellectual tradition (Strydom forthcoming). Although having become quite differently evaluated in the modern period, Aristotle’s distinction between ‘potential infinity’ and ‘actual infinity’ (2015: Book III, Part 6), together with the complementary concept of ‘limit’ (1961: Book Delta, Chapter 17), turned out to be of tradition-founding significance. For him, potential infinity was a never-ending continuous series, progressing either by the addition of another unit or by making another subdivision, without ever reaching its end. By contrast with the continuous nature of potential infinity, he described actual infinity as ‘the infinite in the full sense’, as being ‘complete (teleion), having an ‘end (telos)’, where ‘the end is a limit’. By limit he understood ‘the last point of anything, i.e. the point beyond which it is impossible to find any part of it, but within which all of its parts are found’. Since applied to infinity it meant that actual infinity would amount to all the elements of an adding sequence or of a subdividing sequence being available in their completeness or totality, however, Aristotle summarily rejected actual infinity as a logical impossibility. After the revival of learning, by contrast, no longer beholden to the ancient Greek concern with the concrete and hence their horror infiniti nor the Medieval sequestration of the problem, modern scientists, philosophers and mathematicians both positively embraced infinity and accepted actual or complete infinity as logically perfectly possible (Körner 1968: 30).

In the course of this tradition’s unfolding, Aristotle’s contrast of potential and actual infinity was extrapolated into a fundamental distinction which became stabilized and legitimated, both mathematically and philosophically, and subsequently came to inform many a discipline. In mathematics, it is known as the distinction between the ‘convergent series’ and ‘divergent series’ (Dantzig, 2007: 150-51; Kline, 1990: 1110) and in philosophy Leibniz (1965: 235-36), under the influence of Galileo, introduced the conceptual pair of ‘truths of fact’ and ‘truths of reason’ which proved decisive in Kant’s case. Corresponding to Aristotle’s ‘potential infinity’, mathematics offers pi (π) as the paradigmatic example of a convergent series and its limit. When π is algorithmically derived, a number is obtained which becomes ever longer in decimal places and tends ever closer toward its conventionally accepted value of 3.14159, but the decrease in the distance from that value is infinite while the finite yet ideal limit in principle remains just out of reach. The computer has enabled us to surpass the 700 decimal places added by the nineteenth century, but even with this powerful technology the calculation of π will never finish. By contrast with a convergent series, a divergent series mathematically does not tend toward a finite yet ideal limit like π, but is one in which the addition of more and more units or whole numbers leads to a total that outgrows any boundary one might try to impose. This exponentially accumulating series with its projected sought after totality, its infinite ideal limit, indeed corresponds to Aristotle’s ‘actual’ or ‘complete infinity’ but, contrary to his stance, has since early modernity been taken up enthusiastically and explored with intent.
In philosophy, Kant (1968: B386=A330-B388=A332) took cues from this mathematical distinction via its earlier appropriation by Galileo, Wallis, Newton, Berkeley and Leibniz in order to identify two continuous or infinite series – what he respectively called ‘the descending series…on the side of the conditioned…[representing]…a process of becoming…[which is]…potential’; and ‘the ascending series…on the side of the conditions…already presupposed or given in its completeness’. Of the latter Kant (1968: A332=B389, 1972: 93) held that although we can never comprehend it in its totality, it must nevertheless be enveloped by such a totality which we, consequently, must envisage as a totality given in its completeness if we were to be able to make inferences and judgements. It is this demand for the recognition of an infinite ideal limit to the accumulative divergent or ascending series that led Kant to introduce his characteristic transfinite or transcendental notion of ‘idea or concept of reason’ (Kant 1968: A320=B377) as a logically possible and necessary idealization that captures complete infinity. Such ideas or ‘principles [inherent] in reason’ (Kant 1972: 189) play a key role in all three Kant’s Critiques, in each case appropriately modulated. Although they do not apply to any experiential objects, subjective intuition-based imaginings and morally relevant action and practices, such ideas or principles are nevertheless more than merely useful since, as long as they are internally coherent and consistent, they are actually necessary for the identification, amplification and explication of such objects, aesthetic projections and practical phenomena. More generally, from a cognitive perspective, such a complete totality of conditions refers to our necessary and unavoidable presuppositions without which it is impossible not just to investigate any object domain, but also to cultivate ethical subjects and create and organize society. Most immediately, it is available as the conceptual conditions of social life.\footnote{It is the case, of course, that the conditions of social life include also systems of principles other than language, for example, logic, mathematics and informational redundancy, which are here left by the wayside.}

While the ideas or principles of reason qua infinite ideal limit concepts provide the overall framework of intelligibility and practical and aesthetic orientation relevant to ongoing processes of becoming, this still leaves the finite yet also ideal limit concepts associated with the convergent or descending series – that is, limits comparable to the value of π toward which such processes tend but never reach. For Kant (1968: B383=A327) is concerned not just with the transcendental ideas corresponding to the divergent or ascending series, but also with their ‘immanent…employment’ in structuring the theoretical and practical concepts and aesthetic projections corresponding to the convergent or descending series which give effect to them at the lower level. In this respect, he submits that reason involves not exclusively ideas, but also ‘ideals’, ‘archetypes’, ‘models’ or ‘examples’ (1968: A569-571=B597-599). Ideas have the task of ordering and unifying concepts ‘with a view to obtaining totality in various series’, whereas ideals depend on concepts of the understanding, moral concepts and aesthetic sensibility which in turn unify the elements of objects actions and feelings in a way that allows ‘such series of conditions [to] come into being’ (Kant 1968: A643=B672). As with π as the limit value of the mathematical convergent series, Kant stresses that any attempt to realize the ideal contained in a model or example, whether by knowledge production, action or experience, is ‘impracticable’ (1968: A570=B598) – and that, in the sense that while ‘a great goal…is set before it…it can never of itself reach’ it (Kant 1956: 152). Knowledge, action and experience are always incomplete and perfectly emulating the completeness of the ideal of, say, ‘the wise man’ is an ‘illusion’.

In the course of subsequent post-Kantian developments, the distinction between the convergent or descending and the divergent or ascending series became consolidated thanks to the contributions of both philosophers and mathematicians. Among the former were Maimon, Fichte, Hegel, Comte and Marx and among the latter Cauchy, Weierstrass, Schröder, Dedekind and Cantor, while Peirce played both roles. The critical period in mathematics proved decisive for the stabilization and legitimation of the distinction. Cauchy inaugurated it in the early nineteenth century when he demonstrated in precise mathematical terms that the geometrical sequence can indeed be classified...
into a convergent and a divergent series (Kline 1990: 1110; Dantzig 2007: 150). Later in the century, the inspiring impact of Peirce’s (1992, 1998) logical analyses showed in the coming to a head of the problem of infinity and limits in the seemingly conflicting yet nevertheless complementary contributions of Cantor and Dedekind (Dantzig 2007: 181-82). Together, their contributions in effect clarified the relation between infinite processes and their limit concepts. For the purposes of his astounding set-theoretical mathematics, Cantor adopted a dynamic theory according to which the limit value is generated by the motion of a point on a continuous or infinite line, while Dedekind adopted a static theory according to which the power of the mind by means of a cut (Schnitt) imposes a special classificatory scheme on an infinite process to generate its limit value.

For present purposes, Peirce with a foot in both philosophy and mathematics is of special interest. On inspection, it becomes clear that his accounts of truth and reality were conducted from the viewpoint of both the convergent and divergent series as exemplified by intellectual and scientific investigation. He acknowledged Aristotle’s distinction between potential infinity which is carried by activity (ένεργεία), what he himself called esse in futuro (1998: 180), and actual infinity or complete reality (έντελεχεία). And in keeping with his Kant-inspired ‘pragmaticism’ which focuses on the relation between action and its ‘ends’ as contained in ‘the ideas of human life’ (Peirce 1998: 360, 457, 197) after having critically revised Kant’s (1968: A255=B311, A29, B66=A49, A228) most controversial ‘limit concept’ (Grenzbegriff), the ‘unknowable...thing-in-itself’, he defended actual or complete infinity against the predominant modern nominalist trend under the characteristic title of ‘Thirdness’. According to the convergence viewpoint, ‘[r]eality can be regarded as the limit of the endless series of symbols’ (1998: 323), while ‘[t]he opinion which is fated to be ultimately agreed to by all who investigate, is what we mean by truth’ (1992: 139). As regards divergence, ‘Truth’ is one of the ‘Ends’ (1998: 197) of inquiry and as such a ‘real general’ or mighty force influencing human conduct, which unmistakeably indicates that he understood it as actual or complete infinity. In so far as such an end is a real general, it should be obvious that he conceives reality in accordance with the divergent series. All in all, then, this means that Peirce not only proceeded from the well-established distinction between the convergent and divergent series, but that he simultaneously also identified two corresponding types of limit concepts. For example, truth is on the one hand the ‘finite’ ideal ‘value’ or ‘definite limit’ that the ratio of frequency of investigation has ‘in the long run’ and toward which the ongoing process ‘indefinitely converge[s]’ (1998: 100); on the other, ‘Truth’ as the end of investigation and the entelechy of reality is the infinite ideal limit that lays down the parameters of the former.

III Intermediate systematic reflections

If one now shifts the emphasis to the implications of the reconstructed mathematical-philosophical tradition for the way in which limit concepts could be understood, a strategy of considering the conceptualization of the all-embracing human form of life recommend itself. Starting from the distinction between the convergent and divergent axes, it emerges that it can be conceptualized as consisting of three sets of infinite processes – the objective, the sociocultural and the subjective. Each of these consists of a convergent and a divergent process, that is, an ongoing historical-constructive process of becoming and a structurally accumulating process. The crucial point in the present context, needless to say, is that these two processes have their own particular limit concepts – a finite ideal limit concept and an infinite ideal limit concept respectively – which now apply to all three of the dimensions of the human form of life. Figure 1 below summarizes the results of this reconceptualization plus a few additions for further suggestive clarification. Particularly important among these are the distinctions correlating with the basic one between finite ideal and infinite ideal limit concepts. The cognitively inspired distinction between cultural models and the cognitive order is theoretically decisive, while those between ‘meaning and validity’ (Habermas 2003: 76) and between semantic-pragmatic-symbolic generality and cognitive universality are essential for adequately grasping the specificity of these two types of limit concepts.
Beginning with the convergent axis (a in Figure 1), the objective set covers the ongoing natural-historical process of which it is today in the age of the Anthropocene widely acknowledged that humans form a part and that their form of life is a continuation and extension in a specifically sociocultural form. Second, the sociocultural set houses the ongoing process of the action-based historical creation, construction, production, formation and organization of the sociocultural form of life internally and in relation to nature. Finally, the subjective set consists of the ongoing process of self-cultivation and subject-formation, irrespective of whether individual or social and irrespective of whether under objective and/or sociocultural pressure or, alternatively, autonomously initiating change.

Figure 1: Infinite processes and limit concepts of the human form of life

As regards the divergent axis (b in Figure 1), the three ongoing processes for their part are structurally overdetermined by the concomitant complementary processes of the long-term accumulation and evolutionary stabilization of rational potentials that take on an incursive structuring role in respect of the ongoing processes and, thereby, also enable their recursive regulation. It is these two distinct functions that call for a consideration of the important question of limit concepts. Giving a social-theoretical twist to the mathematical-philosophical notions of infinite ideal and finite ideal limit concepts leads to their conceptualization as cognitive order principles and

---

7 The conceptualization of this dimension is intentionally geared toward admitting what Habermas (2003, 2007) conceives as a weak-naturalistic ontology. The objective world of all possible objects of experience and reference embraces neither just nature as seen by the natural sciences nor just sociocultural reality, but both of these as well as ϕύσις or natura naturans in the sense of the infinite process that spontaneously brings forth a plethora of diverse forms, including humans, and stamps them with a general resemblance to their source.
cultural models on the divergent and convergent axes respectively. Such cognitive order principles as truth, rightness and truthfulness are superordinate in so far as they are structurally available validity concepts occupying the meta-level that exert an incursive indexing, structuring and context-setting force, thus laying down also the parameters of the cultural models. In turn, the cultural models are semantically and symbolically rich constructs that immanently serve as pragmatic directing and guiding or regulative magnitudes toward which action and practices indefinitely tend yet, like π, defy attainment or realization. The selection of the finite ideal limit concepts or cultural models presented in Figure 1 has been made from amongst the vanishing number of possible ones with the intention of reflecting those that have only very recently emerged as, perhaps, the most urgently required directing and guiding orientation complexes under the currently rather challenging conditions.⁸

The extrapolation of the systematic sense of the reconstructed mathematical-philosophical tradition brings to a close the undertaking to recover the hazy background of Habermas’ notion of limit concept. The scene has thus been set to return to him in order to investigate more closely a selection of relevant instances of his understanding and use of the notion from his late theoretical philosophy as presented in Truth and Justification (2003). Throughout, it has to be born in mind that Habermas’ treatment of the notion of limit concept is of central importance to his innovative inauguration of a cognitive turn in Critical Theory. For present purposes, however, the focus is trained on identifying and analysing instances of his treatment that contribute to the sense of a problem regarding limit concepts in his thought. Ultimately, the aim of this friendly critical probe is to bring out a vital aspect of his cognitive departure and, building on it, to contribute to the clarification, consolidation and development of this indispensable dimension of Critical Theory.

IV The problem of limit concepts
For the purposes of the analysis conducted in this final section, two instances of Habermas’ use of the notion of limit concept are selected for discussion. The first is his revision of his long-held concept of truth and his criticism of Peirce’s in line with this rethinking of his position. The second is his depiction of the role of limit concepts in discursive justification in theoretical discourse, particularly truth and ideal or rational acceptability and the relation between them.⁹

Revision of the concept of truth and criticism of Peirce
The first and best access to Habermas’ conception and use of the notion of limit concept is his remarkable revision in the late-1990s of the concept of truth he had held since the late-1960s and early-1970s. The advantage of taking this route is that it allows his decision to be seen not only profiled against the foil of the long-standing debate about truth amongst critical theorists, but also as closely intertwined with his evaluation of Peirce’s conception of truth.

Having instigated a debate about the status of truth in the 1980s, over the next almost two decades Albrecht Wellmer (1986, 1998, 2003) persisted in his criticism of Karl-Otto Apel and Habermas, doggedly targeting their not entirely justifiable assumption of anticipating the realization of the ideal communication community in the real one and thus effectively reaching truth. Instead of a transcendental approach operating with the notion of an ideal, Wellmer (2003) proposed the adoption of a performative version of ‘pragmatism devoid of regulative ideas’. Apel (2003), who

---

⁸ In the discussion below of Habermas’ position on discursive processes aimed at justifying truth claims, he mentions ‘ascertaining truth’, ‘ideal warranted assertibility’, ‘ideal justification’ and ‘rational or ideal acceptability’ that can be considered appropriate cultural models at that level.

⁹ The reasoning developed in these two cases can be applied also to Habermas’ (2003: 237-75) presentation of the difference between truth and rightness. Its inclusion in this essay had originally been envisaged but lack of space put paid to that intention.
never yielded, pointed out that Wellmer proceeds from assumptions not covered by his proposal, while Habermas for years also resisted until the late-1990s when he eventually allowed Wellmer and Cristina Lafont to convince him to rethink his position. The revision followed suit (Habermas 1999: 51, 2003: 38).

Habermas (2003: 37-40) presents his position as him having been compelled to revise his original consensus concept, the ‘discursive conception of rational acceptability’, by relating it to a ‘pragmatically conceived non-epistemic concept of truth’, while simultaneously avoiding the inflation of the former in the form of an ‘epistemic concept of truth’ that cannibalizes the concept of truth as such. This new position is a complex one according to which he actually accepts three different instantiations of the concept of truth – a Kantian, a Platonic and a hybrid one, all of which are different kinds of ‘idealization’ (2003: 99-100). The first is the validity concept of ‘unconditional…[or]…absolute truth’ that serves as the situation-transcendent magnitude to which discursively articulated truth claims necessarily have recourse. The second is the naïve unquestioned pragmatic non-epistemic concept implicitly operative in ordinary everyday speech, action and interaction that subjects in the performative attitude hold as certain and thus free from all imperfections. The third depends on the retention of his original and, indeed, characteristic discursive or procedural view, but transforms his earlier problematic epistemic interpretation which tended to assimilate the unconditional concept of truth to ideal rational assertibility owing to the assumption of a conceptual connection between the two. If one can talk of truth here at all, it is the concept of the ‘truth predicate…[subject to]…cautionary use’ which, more correctly, should be understood as ‘warranted assertibility’ or ‘rational acceptability’. It invokes the epistemically indexed outcome of the discursive treatment of a problematized truth claim which in principle stands under the proviso of ‘fallibility’ (2003: 38) and thus possible future correction.

To explicate this complex revisionist scenario, Habermas embeds his account in a critical evaluation of Peirce’s position on truth. The thrust of the criticism is squarely directed against the American philosopher’s alleged misleading epistemologizing of the concept of truth. It is noteworthy that this argument is possible only on the basis of Habermas’ assumption of a distinction equivalent to that between the convergent axis of the conditioned and the divergent axis of the conditions with their respective limit concepts, but as the criticism unfolds it becomes apparent that he at best operates only implicitly with it so that matters in the end retain a distinct penumbra of opacity. Two relatively lengthy quotations from the same essay in which he traces the genealogical connection between Kant’s ideas of reason and ‘the “idealizing” presuppositions’ of communicative action – a new essay exclusive to the English translation (Habermas 2003: 83-130) – are to the point. According to the first quotation:

‘Peirce wanted to explain “truth” itself epistemically, in terms of progress toward truth. He defined the meaning of truth by anticipating a consensus that all participants in a self-correcting process of inquiry under ideal epistemic conditions would have to attain. The unlimited ideal “community of investigators” constitutes the forum for the “highest court” of reason. There are good reasons against epistemologizing the concept of truth in this way, which assimilates “truth” to “idealized justification” or “ideal warranted assertibility”. Nonetheless, the orientation toward truth – as a property that a proposition “cannot lose” – acquires an indispensable regulative function for fallible processes of justification precisely if

\[10\] Apel must have a point since a ‘pragmatism without regulative ideas’ could not possibly do justice to the complexity of the problem.
such a process can at best lead to decisions about the rational acceptability of propositions and not their truth’ (Habermas 2003: 91-2).

In the second quotation, Habermas reinforces his criticism of Peirce’s epistemologizing of the concept of truth and circumscribes his objection to this move in some more detail:

‘The epistemic reflection of unconditionality is the ideal inflation of the critical audience into a “final” court of appeal. Peirce uses the image of the socially and historically unlimited ideal community of inquirers that continues to pursue the process of inquiry – until they reach the ideal limit of a “final opinion”.

‘This image is misleading in two respects: To begin with, it suggests that truth can be conceived as idealized warranted assertibility, which in turn is assessed in terms of a consensus attained under ideal conditions. But 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. Moreover, Peirce’s image does not direct our attention to the process of justification in the course of which true propositions have to stand up to objections, but to the final state of an agreement not subject to revision. This is contrary to a fallibilist self-understanding that expresses itself in the “cautionary use” of the truth predicate...Despite these objections to an epistemic conception of truth and even after abandoning foundationalist justifications, the idea of a process of argumentation that is as inclusive as possible and that can be continued at any time has an important role in explaining “rational acceptability”, if not “truth”. As fallible, situated beings, we have no other way to ascertain truth than through discourses that are both rational and open-ended.

‘No matter how misleading the image of an ideally extended communication community (Apel) that reaches a warranted consensus under ideal epistemic conditions (Putnam), before an ideal audience (Perelman), or in an ideal speech situation (Habermas), we can in no way forego making some such “idealizations”’ (Habermas 2003: 101).

The two quotations share the same structure. First, Habermas gives a description of Peirce’s position which amounts to epistemologizing the concept of truth. His focus is on the American’s concern with the self-correcting process of the search for truth carried by a community of investigators who discursively exchange counterarguments while being oriented toward the ‘ideal limit of a “final opinion”’. Raising an objection to this strategy, second, he offers a diagnosis of Peirce’s error which, according to him, consists of assimilating unconditional truth to an agreement or consensus among the investigators and of over-emphasizing this final state attained by the community to the detriment of the process of discursive justification. Third, he nevertheless qualifies his criticism by acknowledging that in a certain respect Peirce does have a point. What this qualification is intended to establish is that the image of a self-correcting discursive process oriented toward an ideal limit has to be retained since this is the only way in which humans are able to ascertain the acceptability of a truth claim, but that such fallible acceptance should not be conflated with unconditional truth.

11 The immediately following commentary is focused on Habermas’ criticism of Peirce, with a discussion of his misalignment of the orientation to unconditioned truth and the regulative function reserved for later analysis – see pp. 11-12.

12 Elsewhere Habermas (1998: 360) gives a description in general terms rather than with reference to Peirce of the ‘epistemic...approach as...one which...idealize[s] the justificatory conditions...[and thus]...inflates the idea of a justified assertion to such an extent that truth becomes the limit concept of the justificatory process’, rather than idealized warranted assertibility.
It is remarkable that Habermas traces the source of his own earlier error of reducing unconditional truth to consensus which his self-critical and self-correcting revision is supposed to overcome to nobody less than Peirce. That this is not convincing, however, is borne out by the American’s rather more complex position. It is more probable that the mistake lies with Apel and Habermas himself and, perhaps, also with authors like Putnam and Perelman of whom they took note in the 1960s. As pointed out earlier, Peirce the mathematician-philosopher not only started from the well-established distinction between potential infinity and actual or complete infinity and, by extension, between the convergent and divergent series, but he simultaneously also identified two quite different corresponding types of limit concepts. In accordance with the convergent series, on the one hand, truth for him is the ‘finite...[ideal]...value’ or ‘definite limit’ that the ratio of frequency of investigation has ‘in the long run’ and toward which the ongoing process is most immediately oriented and toward which it ‘indefinitely converge[s]’ (Peirce 1998: 100) but, like π, can only approximate and never reach. In accordance with the divergent series, on the other hand, the ‘absolute truth for all questions’, ‘Truth’ as one of the eternal, universal and necessary ‘Ends’ of human endeavour together with ‘Right and Beauty’ (Peirce 1998: 419, 123, 197), is the infinite ideal limit that represents the conditions of the former.

In Habermas’ interpretation, Peirce is confined to the convergent axis while he effectively makes his own the convergent-divergent distinction which is actually already available in the American. The apparently paradoxical qualification with which he closes his criticism of Peirce, namely that the latter’s ideal model is misleading yet that we nevertheless cannot avoid making such an idealization, is designed not just to rescue his own earlier conception of truth, but also to restrict Peirce to that level. Presenting his own position following the criticism of Peirce, by contrast, he adopts the relational set of ‘finite, spatiotemporally limited discourses’ together with their ‘ideal model of a ‘continuous...[or] ...“endless conversation”’ and the concept of ‘absolutely valid truth’ (Habermas 2003: 102). When, continuing, he identifies truth as the ideal magnitude toward which ‘the participants are oriented’ and which, therefore, ‘is reflected at the level of the discursive ascertainment of truth in performative idealizations’, he portrays – as should be the case – the emergent evolutionarily stabilized concept of unconditional truth as over-determining the ideal model of the ongoing discursive process and, thereby, also effectively touches on two different types of limit concepts – ‘absolutely valid truth’ and ‘performative presuppositions’. At the same time, however, it is evident that his treatment of these limit concepts leaves something to be desired, despite his obvious assumption of the conditioned-conditions – effectively the convergent-divergent – framework. This conceptual pair is of course an inherent part of Habermas’ thinking, deriving from Kant’s architectonic distinction between understanding and reason, but here we have an unmistakable indication of aspects of the problem of limit concepts.

In general terms, the problem is of the nature of a structural deficit – a deficit that could be corrected, to be sure, were Habermas (2003: 218) to fill in his undeveloped vision of the ‘cultural embodiment of reason’. More specifically, however, the problem concerns a proper distinction between and characterization of the two types of limit concepts as well as the relation between them. First, a limit concept such as ‘truth’, like rightness and truthfulness, which Habermas rightly considers as Kantian totalities, are not just philosophical or, more generally, linguistic concepts or even idealizations, but they are also cultural structures. As articulations of reason, they are situation-transcendent or meta-cultural structures and as such form part of the cognitive order of the human

---

13 Exactly the same pattern is to be found in Peirce: on the one hand, ‘holding for true’ is of two kinds; the one is that practical holding for true which alone is entitled to the name Belief, while the other is that acceptance of a proposition which in the intention of pure science remains always provisional’ and, on the other, ‘the eternal...[ideal] Truth...[which is]...infinite...[and]...generative...(1998: 56, 123).

14 Accordingly, Habermas writes: ‘conditions that must be fulfilled so that what is conditioned can take on...values’ (2003: 86, my emphasis).
sociocultural form of life. By contrast, the ‘ideal model of a ‘continuous...[or]...“endless conversation’” and others such as ‘idealized warranted assertibility’, which again plausibly Habermas considers as Kantian-Platonic hybrids, are immanent cultural models operative in actually ongoing processes of discursive justification. Besides these, various others appear randomly throughout his writings. Most characteristically since the 1970s are pragmatic presuppositions operating as idealizations such as publicity and inclusiveness, equal rights in communication, exclusion of deception and illusion as well as absence of coercion, but there are also ones applicable to theoretical discourse like ideal ascertaining of truth, ideal justification and rational or ideal acceptability. While these are all abstract or strictly cognitive cultural models directing and guiding processes of the search for truth, it is quite possible to identify more substantive semantically and symbolically rich cultural models playing a pragmatic role in particular situational contexts – for example, those relevant to scientific arguments concerning a specific hypothesis.

As cultural structures, whether transcendent or immanent, all these idealizations serve in some sense as orientation complexes. In an attempt to characterize the limit concepts in question, however, it is important to note that the two types of cultural structures quo orientation complexes differ appreciably from one another. In this respect, Habermas’ depiction of participants as being ‘oriented toward truth’ is in need of elaboration. Participants in a discursive process are most immediately and directly oriented toward the immanent cultural models relevant to their particular search for truth, including both abstract cognitive and substantive semantic-pragmatic-symbolic models which are essential for the articulation of the texture and details of the process and their practice. Their orientation to the meta-cultural cognitive order presupposition of truth, by contrast, is of a different order altogether. Orientation in this case is only mediate, indirect and diffuse in so far as this orientation complex at most delineates, indexes and structures the context of the process and practice in which the participants are engaged. For its implementation it requires a range of cultural models. This is what is implied by the phrase the ‘cultural embodiment of reason’.

Second, Habermas brings this question of the relation between the two types of limit concepts to a head when he submits in the quotation above that ‘the concept of an absolutely valid truth is reflected at the level of the discursive ascertainment of truth in performative idealizations’. The word ‘reflected’, like the word ‘oriented’ just discussed, is revealing. It is rather opaque shorthand for the relation between the two limit concepts which covers over the distinct role played by each of them. Transcendent cognitive order presuppositions like truth have an incursive indexing and structuring significance and potency in relation to immanent cultural models, while the latter have a recursively regulative role of specifying, amplifying and moderating the incursive framing effect of the meta-cultural structures in repeated iterations within particular situations. There is a certain degree of equivocation or, at least, lack of clarity observable in Habermas’ writings regarding the seat of regulation which derives from a corresponding absence of a reflective and differentiated use of the notion of limit concept. Most often it appears as though it is truth that exerts a regulative function, yet on one occasion he does write that ‘it is not truth as such but the epistemic concept of ascertaining truth that is the regulative idea guiding our practices of inquiry and justification’ (Habermas 2003: 227). Here it should be pointed out, however, that while it is correct to assign the regulative function to the finite ideal limit concept rather than to the infinite ideal limit concept, it is by no means the case that the ongoing search for truth is directed and guided by only one such cultural model. On the contrary, from situation to situation there is a whole complex of interrelated

---

15 If one takes Habermas’ (2003, 2007) conception of ‘weak naturalism’ seriously, it would have to be acknowledged furthermore that a meta-cultural cognitive order presupposition like truth both has a natural-evolutionary root and a dimension of independent cultural elaboration and articulation. It presupposes the enccephalization and phylogenesis of the brain and the concomitant emergence of the cognitively fluid mind characteristic of *Homo sapiens sapiens* that made possible the cultural or human revolution of circa 30,000 years ago which, in turn, cleared the way for cultural evolution to stabilize the cognitive order beyond nature.
models that fulfils this role. An elaboration of the cultural embodiment of reason, particularly if done in a cognitively inspired mode, would undoubtedly confirm this contention.\footnote{16}

**Discursive justificatory paradox and fallibilistic proviso**

Above it was argued that the core issue regarding the problem of limit concepts in Habermas is a structural deficit that derives from the not entirely adequate treatment of the different ideal cultural structures serving as orientation complexes – and that, despite his insistence on the configuration of the nonepistemic concept of truth, ideal assertibility and unconditional truth within the implicitly accepted framework of the convergence-divergence distinction. That this state of affairs is not without consequences is confirmed by a brief analysis of a revealing passage in which he writes as follows:

‘This nonepistemic concept of truth, which manifests itself only operatively, that is, unthematistically, in action, provides discursively thematized truth claims with a basis for a justification-transcendent point of orientation. It is the goal of justifications to discover a truth that exceeds all justifications. This transcending relation guarantees the difference between truth and rational acceptability, but puts the participants in discourse in a paradoxical position. On the one hand, they are able to vindicate controversial truth claims only thanks to the convincing power of good reasons. On the other hand, even the best reasons are under the proviso of fallibility so that precisely at the point where the truth and falsity of propositions is the only issue, the gap between rational acceptability and truth cannot be bridged’ (Habermas 2003: 39-40, translation of first sentence altered with reference to Habermas 1999: 53).

The key second sentence determining the content of the paragraph stands out as especially problematic. Against the background of the mathematical-philosophical understanding of limit concepts, it is incontrovertible that truth, absolute and unconditional truth, is not and cannot be ‘the goal of justifications’. This statement is obviously misleading in that it perpetuates the conflation of aspects that need to be distinguished carefully and consistently. Such conflation is contrary to Habermas’ own intention and the very thrust of his revision. The cognitive order presupposition or principle of truth is once again brought down to the level of discursive justification and conflated with the cultural model which serves as the goal and, hence, the regulative instance of the process. However, this observation by no means entails the denial of the discursive justificatory process having a transcending relation to truth which goes above and beyond rational acceptability. This relation is established and maintained by a principle-reflexive orientation toward unconditional truth and it is this orientation that secures the ‘property of propositions that they “cannot lose”’, in Habermas’ (2003: 38, 91) phrase.\footnote{17} As such, this orientation differs sharply from the spatiotemporally situated process’ rule-reflexive orientation toward its particular goal of attaining, even if in principle only by approximation, the warranted assertion of a particular proposition, its proper justification, its rational acceptance and thus ascertainment of its in principle fallible truth – the *cautionary use* of the truth predicate’, according to Habermas (2003: 38), being imperative in this case.

\footnote{16}{Footnote 8 above contains a list of what can be considered cultural models relevant to the discursive justification of a truth claim that Habermas mentions.}

\footnote{17}{The cogency of the distinction between rule-reflexivity and principle-reflexivity is confirmed by Habermas’ remarkable phrase: ‘to orient her actions by rules whose concept she has mastered’ (2003: 95, my emphasis) which, of course, correlates with his distinction between ‘meaning and validity’ (2003: 76). All cognitive order principles, including unconditional truth, form part of the conceptual conditions of social life which, in turn, make all rules in the form of cultural models possible.}
The unfortunate conflation in the second sentence sets the scene for the series of misalignments following in the rest of the paragraph. Here it seems as though Habermas bases his argument on the assumption that in a certain respect reasons can be aligned with the transcendent dimension occupied by unconditional truth, but then he submits that such reasons are in principle under the proviso of fallibility, from which for him follows that there is an unbridgeable gap between truth and rational acceptability. While difficult to spot immediately, the first misalignment due to an obscure mediation relation should be obvious. The transcendent cognitive order principles, including truth, do not provide any aspects of the reasons articulated in discursive justification, but by their incursive structuring potency only make available the universal cognitive or conceptual conditions for the construction of appropriate reasons. The reasons themselves occur immanently on the discursive level and their mediated makeup is intelligible only by recourse to the inferential dialectics of embodied abduction, conceptual deduction and empirical induction – to invoke Peirce (1992, 1998) – at that level. The deductive component of a reason does not derive directly from the principle of unconditional truth, but rather from the particular conceptual framework or theory appropriate to the proposition under discussion. Since Habermas’ configuration of action/nonepistemic truth and discourse/warranted acceptability on the convergent axis and unconditional truth on the divergent axis apparently corresponds directly to the inferential pattern of abduction-induction-deduction, he tends to overlook that the inferential dialectics pertinent to the testing of a given proposition plays off fully within the process of discursive justification while only indirectly being conditioned by the transcendent principle of unconditional truth.

The second misalignment concerns Habermas’ depiction of the discursive justificatory paradox and the gap from which it derives as well as the associated fallibilist proviso. At the very point where the truth and falsity of propositions are weighed and decided, according to him, an unbridgeable gap opens between rational acceptability and truth, thus manifesting a paradoxical state that hangs the sword of fallibility over any and every vindicated and accepted knowledge or truth claim. Now, whereas Habermas stresses the gap between rational acceptability and truth, the two limit concepts marking the convergent and divergent axes respectively, it needs to be pointed out here that the paradox derives less from this particular gap, as he submits, than from another that remains unregistered in his account – namely, the gap appearing on the convergent axis of the ongoing process of the justification of a particular proposition. It is the one between the discursively vindicated or justified truth claim and the ideal of rational acceptability that serves as the goal of the process and plays a direct regulative role in the establishment of the truth claim. Since rational acceptability, comparable to the value of π, is a finite ideal limit concept toward which the knowledge production process indefinitely tends yet is unable and never will be able actually to reach, claims to the truth are permanently subject to a fallibilist proviso. Writing on one occasion that ‘the wound opened up in everyday practice by a truth claim that has become problematic must be healed in a discourse that cannot be terminated “once and for all”, Habermas (2003:101) apparently knows this, but he seems not to make the necessary connections. Far from depending on the gap between rational acceptability and truth, the fallibilist proviso derives from the fact that however long discourse, justification, acceptance and thus knowledge production continue, the limit concepts qua cultural models directing and guiding this process of becoming, such as ascertaining truth, ideal warranted assertibility, ideal justification and rational acceptability, rather than admitting of complete realization, keep on receding, always staying a step ahead, like π, and thus permanently occluding full attainment.

Conclusion
The investigation of Habermas’ conception and use of the notion of limit concept, even given the partial nature of this probe due to its focus on truth alone, draws attention to just how significant this concept is – a significance unequivocally confirmed by both mathematical and philosophical thought. As on the one hand a basic constraining classification device or scheme imposed by the
human mind and on the other an ideal value generated by practices, it is necessary not only for the constitution of domains, objects or phenomena in the face of infinite processes marking the continuous underlying reality, but also for making the kind of distinctions required for the organization and regulation of justificatory discourses and, more broadly, of social life itself. Although limit concepts for the most part operate intuitively and unthematically in human activities, barring mathematics and engineering, it is obviously a concept that demands reflection in terms of its essentially cognitive nature.

Both philosophically and social-theoretically, Habermas has done more in recent times than anyone else to make the concept visible, even if not intentionally, particularly by his conception of discursive justification and his highly suggestive formal-pragmatics. These pointers cry out for the explicit treatment of the concept as such and its exploration and development from a cognitive-theoretical perspective relevant to both philosophy and social theory as well as, by extension, to Critical Theory. The various issues indicating a problem of limit concepts in Habermas all contribute to the build-up of pressure toward beginning the task of explicating the notion and linking it to cognitive-theoretical thinking. This was the intention of the analysis presented in this essay, but if its significance and implications for Critical Theory were to be more fully extrapolated, this exercise would have to be expanded beyond the concept of truth so as to include especially rightness but also truthfulness as well as their cultural embodiment in a range of more specific abstract cognitive and substantive semantic-pragmatic-symbolic cultural models operative in social life. This would be one way of fruitfully carrying forward Habermas’ rich and inspiring legacy.

References
Leibniz (1965 [1714]) The Monadology and Other Philosophical Writings. London: Oxford University Press.