



Análisis Filosófico

ISSN: 0326-1301

ISSN: 1851-9636

info@analisisfilosofico.org

Sociedad Argentina de Análisis Filosófico
Argentina

Cormick, Claudio Javier

Habermas, Rorty, and the problem of competent interlocutors

Análisis Filosófico, vol. 40, núm. 2, 2020, Noviembre-, pp. 213-246

Sociedad Argentina de Análisis Filosófico

Buenos Aires, Argentina

DOI: <https://doi.org/10.36446/af.2020.368>

Disponible en: <http://www.redalyc.org/articulo.oa?id=340065365004>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica Redalyc

Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto

HABERMAS, RORTY, AND THE PROBLEM OF COMPETENT INTERLOCUTORS*

Habermas, Rorty y el problema de los interlocutores competentes

CLAUDIO JAVIER CORMICK^a

<https://orcid.org/0000-0003-0162-2429>

ccormick@filo.uba.ar

^a Instituto de Investigaciones Filosóficas - Sociedad Argentina de Análisis Filosófico - Consejo Nacional de Investigaciones Científicas y Técnicas, Buenos Aires, Argentina

Abstract

In texts such as “Richard Rorty’s Pragmatic Turn” Jürgen Habermas defends a theory that associates, *on the one hand*, the truth-claim raised by a speaker for a proposition *p* with, *on the other hand*, the requirement that *p* be “defendable on the basis of good reasons [...] at any time and against anybody”. This, as is known, has been the target of criticisms by Rorty, who—in spite of agreeing with Habermas on the central tenet that the way of evaluating our beliefs must be argumentative practice—declares that the only “ideal presupposed by discourse” is “that of being able to justify your beliefs to a *competent* audience”. We will consider two texts from 1971 –surprisingly neglected in most approaches to the debate– in which Habermas *did* include such a “competence condition” to elucidate the notion of truth. We will discuss whether there are good reasons to relinquish such a condition and to refer, instead, only to the formal or procedural properties of argumentative exchanges, as Habermas does in presenting the notion of “ideal speech situation”. As we will try to argue, there are no such good reasons.

Key words: Richard Rorty; Jürgen Habermas; Universalism; Ethnocentrism.

Resumen

En textos como “El giro pragmático de Richard Rorty”, Jürgen Habermas defiende una teoría que asocia, *por un lado*, la pretensión de verdad planteada por un hablante para una proposición *p* con, *por otro lado*, el requisito de que *p* sea “defendible sobre la base de buenas razones [...] en cualquier momento y frente a cualquiera”. Esto, como

*I wish to thank an anonymous referee for *Análisis Filosófico* for valuable remarks about a previous version of this article. I would also like to thank my colleagues in the pragmatism group at SADAF (Santiago Armando, Jonathan Erenfry, Anahí Grenikoff, Pedro Martínez Romagosa, Bruno Muntaabski, Daniel Pared, Federico Penelas, Alejandro Petrone and Mauro Santelli) to whom this work was first presented.

se sabe, ha sido el blanco de críticas de Rorty, quien (pese a compartir con Habermas la tesis central de que la práctica argumentativa ha de ser la forma de evaluar nuestras creencias) declara que el único “ideal presupuesto por el discurso” es “el de poder justificar las propias creencias frente a un público *competente*”. Consideraremos dos textos de 1971 –sorprendentemente omitidos en la mayor parte de los acercamientos al debate– en los cuales Habermas de hecho incluyó, para analizar la noción de verdad, una “condición de competencia” como la mencionada, y discutiremos si existen buenas razones para abandonar una condición tal y referirse, en su lugar, únicamente a las propiedades formales o procedimentales de los intercambios argumentativos, como lo hace Habermas al presentar la noción de “situación ideal de habla”. Como intentaremos argumentar, no hay tales buenas razones.

Palabras clave: Richard Rorty; Jürgen Habermas; Universalismo; Etnocentrismo.

1. Introduction and outline

“I agree with Apel and Habermas”, Richard Rorty has written, “that Peirce was right in telling us to talk about discourse rather than about consciousness, but I think that the *only* ideal presupposed by discourse is that of being able to justify your beliefs to a *competent* audience”. And, he adds, “everything depends upon what constitutes a competent audience” (Rorty, 2000, p. 9. Emphasis in the original). In contrast to this “ideal”, Rorty is criticizing what he takes to be the Habermasian tenet that the ideal presupposed by discourse is that of being able to justify one’s beliefs *to anyone*—an elucidation that includes no “competence condition”. Now, quite surprisingly, there are two texts, not mentioned by Rorty, where Habermas *does* include such a competence clause: Habermas’ Gauss Lectures of 1971, and parallel passages in his *Theorie der Gesellschaft oder Sozialtechnologie?*, of the same year (Habermas, 2002; Habermas & Luhmann, 1971/1979).

We will try to prove two points. As a purely historical matter, we will show, *contra* Rorty, that Habermas did, once, make room for a competence condition in his theory. Therefore, the right question to tackle when considering the differences between Habermas and Rorty should not be “why does Habermas not include a competence condition like the one Rorty demands?” but, in any case, “why did Habermas *relinquish* such a notion, that had been included in earlier versions of his theory?”. We will try to show that ignoring the early Habermasian references to the problem of competence *obscures* what his differences with Rorty consist of. As a strictly philosophical, not only historical, point, we will try to show that Habermas’ abandonment of the competence condition is not well justified. In other words, that his own references

to the role of the notion of “competent subject” for the elucidation of “truth” or “justification” do not provide us with good grounds to move, as Habermas does in later texts, to an “unqualified” universalism that appeals to the notion of “rational consensus” *without* specifying that consensuses with a probative value must be those among competent interlocutors.

We will proceed as follows: in section 2.1, we will briefly revise Rorty’s counterexamples to Habermas’ universalism—his criticism that such a theory entails a demand to justify one’s beliefs before Attila, a Nazi or a three-year-old. In section 2.2, in order not to take Rorty’s examples at face value, we will assess them by exploring, first, the *scope* of Habermas’ proposal (namely, if it in fact refers to discussing our views against literally everyone) and, second, the *character* of such proposal (namely, if it is some kind of epistemic deontology). We will need to distinguish between our obligations as “spreaders of the word” and as “truth-seekers”, insofar as it is only the second possibility that can be supported by Habermas’ claims about entering in symmetrical argumentative exchanges (which *do not* imply that we have any obligation to *teach* children, or Nazis). To show more clearly what kind of proposal Habermas’ is, in section 2.3 we will introduce the German author’s *ipsissima uerba* in some texts from the early 1970s, in which he presents an attempt to *define* truth in terms of universal agreement under certain circumstances. With this attempted definition in mind, we will return, in section 2.4, to Rorty’s counterexamples in order to assess whether this early Habermasian proposal actually makes it reasonable to demand us to defend our beliefs against every possible interlocutor. We will consider Wellmer’s remarks that the rational obligation not to “protect” our beliefs against the possibility of refutation *does not entail* a rational obligation to enter into every possible argumentative exchange. There can be certain propositions which we can reasonably discard as simply not worth discussing – and also, which is precisely Rorty’s point, certain subjects that we may discard, as a general rule, as simply incapable of offering good objections against our beliefs.

Now, given that Habermas’ early attempt to provide an epistemic definition of truth is an element that is later abandoned, we will need to consider, in section 2.5, the turn Habermas takes in 1996, when he turns, instead, to an elucidation of the problem of what it is to “convince ourselves” of the truth of a proposition – that is, the turn from a theory of truth to a theory of justification. We will point out, in section 2.6, that the key difficulty pointed out by Rorty affects this new version of the theory as well: even if the notion of rational consensus is not

used to define “truth” but to elucidate “justification”, the problem is still to determine why such consensus should not be restricted only to *competent* interlocutors.

Now, as we anticipated, our strictly historical point is that Habermas did once consider the competence condition—and later relinquished it; to prove this, we will tackle more deeply, in section 3, Habermas’ two texts from 1971. Habermas’ 1971 proposal, however, does not strictly coincide with what Rorty will demand. According to Habermas, whereas the notion of “truth” is indeed dependent on that of rational agreement between competent subjects, the condition of “competent subject” can *itself* only be attributed on the basis of a rational consensus—and therefore is not susceptible to being used as an “independent” notion, conceptually prior to the problem of rational agreement. In Habermas’ view, then, a circle emerges between, on the one hand, the problem of determining whether certain *de facto* agreement is indeed relevant for the establishment of the truth of a proposition, and, on the other hand, the problem of determining whether or not certain subjects are in fact competent.

This is, as we will see in section 4, a circle that Habermas wishes to overcome by means of the notion of “ideal speech situation”. However, as we will try to show, the introduction of this notion does not *replace* the need for a competence condition. Rather, we will argue – on the basis of a discussion in Habermas’ “Wahrheitstheorien”, originally published in 1973 – that the kind of procedural aspects that characterize an ideal speech situation cannot work a *creatio ex nihilo* of good arguments.

Finally, in section 5, we will try to show, *via* a detour through an exchange of criticisms between Rorty and Albrecht Wellmer, that the introduction of the notion of “competent subject” raises the obvious problem of non-neutrality between communities holding different beliefs: the status of “competent subject” will not be attributed in one of such communities to the same individuals as in the other. “Competent” will then mean “competent *for us*”. Nevertheless, we will hold that this should not raise a special difficulty for Habermas, insofar as his own epistemology of the social sciences is distinguished by the acknowledgment that practitioners of such sciences can be methodologically compelled to impose *their own* standards of rationality, if necessary, against those of the subjects they are trying to understand. Habermas’ theory, from this point of view, *already* needs to make room for a non-neutral appeal to the standards of one’s culture.

2. Rorty's objection and some preliminary clarifications

2.1. *Attila, the Nazi, the three-year-old, and traditional Chinese physicians: Rorty's criticism of Habermas' universalism*

As is known, the universalism that Rorty finds -and criticizes- in Habermas consists in the idea that, in order to analyze the truth claims that we raise, we need to appeal to some notion of “*universal* community of justification”. In “Universality and truth”, the American neo-pragmatist questions such a universalistic proposal by presenting a series of obvious counter-examples, such as the implausibility of trying to defend a belief about justice to Attila, or about trigonometry to three-year-olds (Rorty, 2000, p. 11), and assumes that universalistic philosophers might actually mean, not that they are in fact engaged in such desperate attempts, but that they want to *educate* their interlocutors (not only children but also, for example, Nazis) so that they become able to accept more reasonable beliefs (Rorty, 2000, p. 19). In any case, according to Rorty, universalist philosophers should become more aware that human beings are split into different communities of justification on the basis of a sufficient overlapping between their respective beliefs and desires (Rorty, 2000, p. 15). “None of us”, writes Rorty, “take all audiences seriously; we all reject requests for justification from some audiences as a waste of time. (Consider the surgeon refusing to justify her procedure to Christian Scientists, or to Chinese physicians who suggest relying on acupuncture and moxibustion)” (Rorty, 2000, p. 27). In this way, Rorty is rephrasing his previous claim that “everybody is ethnocentric when engaged in actual debate”, because to be one is simply “to divide the human race into the people to whom one must justify one's beliefs and the others. The first group—one's *ethnos*—comprises those who share enough of one's beliefs to make fruitful conversation possible” (Rorty, 1984b, p. 12).

2.2. *Are Rorty's counterexamples good?*

There are two points that seem to be distinguishable in Habermas' universalism and which raise Rorty's doubts. The first point refers to the *scope* of the community of justification which Habermas might be theorizing about (are we literally speaking about *every* language user?); the second, to the *specific kind of proposal* that Habermas seems to present in his discourse theory (is it an *epistemic deontology*?).

Let us consider the question of the scope. The very title of Rorty's “Universality and truth” refers to a connection that is, according

to Rorty, rather doubtful, and, according to Habermas, some kind of platitude, an “intuition” which is not really in question but simply requires clarification. So, instead of taking Rorty’s counterexamples at face value, let us try to determine what connection a discourse theory may be entitled (or supposed) to establish between the problem of truth and that of a *universal* community of justification. Habermas, as is known, refers to defending our truth claims “in all possible contexts, that is, at any time *and against anybody*”. When declaring this, Habermas is attempting to express “the intuition [...] that true propositions are resistant to spatially, socially, and temporally unconstrained attempts to refute them” (Habermas, 2000, p. 46). From the point of view of this passage, the connection between universality and truth consists in the following: a proposition cannot be true if it only resists the attempts to refute it *made by some people*; if it is actually true, it must resist the attempts to refute it *made by anybody*. In other words, it would be odd to say something like “I believe *p* is true, so I expect it to resist the attempts to refute it made by subjects S1, S2 and S3, but I do not expect it to resist the attempts made by subject S4”. Now, why is this? The obvious reason to reject this last analysis of *p*’s truth is that, when it comes to the truth-value of *p*, it should not matter *by whom* an objection to *p* is raised in dialogue. If *p* is actually true, it will “resist” attempts of refutation – that is to say, more simply, *it will not be refuted*, period. Claiming that its being true means that it will not be refuted *by a specific set of individuals*, while leaving open the possibility of its being refuted by *other* individuals, is as irrational as specifying that it will resist objections in French, but conceding that it might not resist objections in German. Now, how does this analysis leave us concerning Rorty’s counterexamples?

Whereas a first reading of the debate – according to which the existence of three-year-olds becomes a knockdown evidence against a discourse theory of truth – risked being unfair to Habermas, we now risk – after analyzing the Habermasian articulation between truth and universality by means of the notion of “resistance to attempts of refutation” – being unfair *to Rorty*. If Habermas is simply stating a platitude such as “a true proposition cannot be refuted *by anyone*”, which follows from the fact that we *define* “refutation” as “proof of the falsity of a proposition”, how can Rorty disagree with *that*?

Therefore, if we exclude the possibility of Rorty disagreeing with a platitude, his objections must be aiming at something which is *not* the trivial tenet concerning the “resistance” of true propositions. If that is the case, then precisely the platitudinous character of such a tenet prevents it from being relevant to the question at hand.

This leads us to the second problem that this section needs to address. Namely, whether Habermas is presenting something like an epistemic deontology, *i.e.*, stating what we, as epistemic agents, are obliged to do –as different from reconstructing the conditions of a statement's *being* true, regardless of any obligation an agent may have.¹

This deontology, in fact, seems to be the aim of Rorty's attacks when the latter refers to possible *activities to engage in* (such as *defending* one's views to three-year-olds); if this is the case, then what is under attack is not a platitude anymore. Quite on the contrary, we will have moved from the uncontroversial tenet "a true proposition cannot be refuted by anyone" to the more troublesome maxim "if we believe that a proposition is true, we should engage in the activity of defending it to anyone". This maxim does not follow from the platitude about the "resistance" exerted by true propositions – in fact, the platitude in question does not refer to anything anybody should do or refrain from doing; it only attempts to elucidate the meaning of "true", as an alternative to an elucidation such as "a true proposition *corresponds to the states of affairs*", and is as little "deontological" as correspondence theory. We must add something to Habermas' platitude if an epistemic obligation is to emerge. And, after adding that extra element, Habermas' point might become less plausible against Rorty's objections.

The element needs to be something that appropriately connects, on the one hand, the conditions for the truth of a proposition and, on the other, a series of obligations. How can it be that, if those conditions are elucidated in terms of "resistance to objections", an obligation emerges to defend those propositions that we believe are true "against anybody" –including such scarcely promissory interlocutors as Nazis, Attila, and three-year-olds?

A possible line of defense of Habermas' position, understood as a deontology, consists in declaring that the contingent, empirical obstacles for fulfilling a certain obligation do not cancel such an obligation².

¹ Kyung Man Kim, for example, takes it for granted that Habermas is literally claiming that we have an obligation to defend our beliefs against *anyone*. This is why Kim appeals to the alleged fact that even subjects paradigmatically considered rational, as natural scientists, do not fulfill the standards of a Habermasian communicative conception of reason, insofar as they "do not think that the 'tacit knowledge' they use [...] has to be justified to people like Habermas who are outside of the scientific community. Unlike Habermas, they believe that it is pointless to justify scientific rationality to those outside of science" (Kim, 2014, p. 111, cf. also 2011, pp. 398, 402).

² This is an objection that was raised by an anonymous referee for *Análisis Filosófico*, whom I thank for helping me make the point clearer.

According to this line of defense, it is not irrational to claim that *in principle* we should try to defend our beliefs before even the kinds of subjects Rorty refers to in his counterexamples; the universalist can think that, given enough time, he or she will be able to convince even the Nazi and, that, as a consequence, it is not irrational to *try* to do it. Of course, the defense will continue to say, this does not mean that we can expect to convince the Nazi in the short term, but of course we can have lots of obligations that cannot be fulfilled *immediately*.

Now, this possible reply to Rorty's objections assumes (1) that we have an epistemic obligation, according to universalism, to *try to convince* anyone of our own beliefs and (2) that (on the basis of the first assumption) the only reason why a critic may deny the existence of this obligation is that we cannot have a duty to attempt the *impossible*. Consequently, if convincing Nazis of certain moral beliefs is not impossible *in principle*, but only more difficult than convincing other kinds of subjects, then the universalistic duty of defending our views *to literally anyone* remains.

Now, is it reasonable to assume that we have a "pedagogic" obligation, such as the one described in (1)? Well, it *might* be: particularly in what concerns *moral* beliefs, it seems reasonable to assume that if "fighting Nazism" is a *moral* obligation, then "converting Nazis" must be an obligation as well—we human beings act in the ways we do because of, *inter alia*, the kinds of beliefs we have, so it would be irrational to say that we have a moral obligation to, say, physically fight Nazis but *not* an obligation to discuss their views. However, even if this kind of example can support the idea that we have a "pedagogic" obligation in the moral realm, it would not be enough to counter Rorty's criticisms, simply because it cannot be applied to theoretical beliefs, such as the case of geometry and children. Do we have a *duty* to teach children everything we believe, no matter how abstract? Why would we?

I do not believe, then, that the distinction between obligations "in principle" and the empirical circumstances in which those obligations are to be fulfilled is a good reply to Rorty. In any case, what really matters is that Habermas does not *need* to vindicate the sort of "pedagogic" obligation described in (1). In fact, the whole scenario we are describing, in which an agent enters in a communicative exchange with another one with the purpose of *teaching* him or her (that is, an *asymmetric* exchange) is alien to Habermas' proposal, as Rorty well knows. If "universality" and "truth" are connected by discourse theory under the form of a deontology, this cannot be because of an obligation to "spread the word" but (if there is a deontology at all) because of an

obligation to *seek* the truth. Habermas' discourse theory does not refer at all to "pedagogic" obligations; it simply attempts to elucidate (in early formulations) what *truth* is and (in texts from 1996 onwards) what it takes for a belief to be *justified*. Given that the status of a proposition will have to be determined on the basis of how it "fares" in argumentative exchanges, we are obliged, as truth-seekers, to expose that proposition to as many objections as possible – a kind of demand that is sensitive to the specific *discursive* aspect of Habermas' proposal; the idea must be then that Habermas' *discourse* theory of truth entails an epistemic deontology which differs from that which would follow from, say, a theory according to which the criterion of truth was the result of observations carried out by an isolated subject.

Now, if this obligation (as, recall, "truth-seekers", *not* as "spreaders of the word") actually has a strictly universalistic form (i.e., if it entails that we need to *try to justify* our beliefs against *literally anyone*), it must be supported by the prior tenet that the "probative value" of a discursive exchange somehow depends on the unrestricted nature of a community of justification. In other words, in order to support the deontological claim that, as truth-seekers, we are obliged to defend our views against anybody, and not only against those we take to be *competent* interlocutors, the claim must also be made that the *truth* of a proposition *p* has to be defined by reference to the possible agreement of literally everyone – or, in a weaker version of the theory, that the *establishment* of the truth of a proposition *p* has to be defined in this way.

The point is that searching for an epistemic deontology in Habermas' proposal must, in any case, be an *indirect* task: we need to turn now to his discourse theory in order to determine whether a reasonable connection with strictly universalistic conditions emerges. Before returning to Rorty's counterexamples, we must, then, take a detour through Habermas' *ipsissima uerba*.

2.3. What is Habermas actually doing? Habermas' theory of truth in the early 1970s

In light of the ethnocentric criticisms of Habermas' universalism, it seems relevant to try to shed some light on what his consensus theory amounts to. To determine what our obligations as truth-seekers are, in the context of Habermas' proposal, we need to distinguish between the following possibilities:

A. Habermas is trying to elucidate in universalistic terms *what it is for a proposition to be true*. That is to say, his universalism claims

that there is a connection (in terms of necessary conditions, of sufficient conditions, or both) between a proposition being true and the proposition being universally agreed upon under certain circumstances.

B. Habermas is trying to elucidate in universalistic terms *the strictly epistemic fact that a proposition being true has been adequately established* (which has to be distinguished from the “fact”, not epistemic itself, that a proposition *is* actually true). That is, his universalism claims that there is a connection between the universal agreement on a proposition and its being adequately established as true.

Option A is certainly a reading that fits the letter of Habermas’ discourse theory of truth as it was presented in the early 1970’s, as an explicit alternative to what Habermas calls an “ontological” theory of truth—namely, correspondence theory. Whereas the latter would fall prey to the circumstance that it is not possible to define truth in terms of correspondence with reality, because “reality” is itself a notion that can only refer to the correlate of *true sentences* (that is to say: “reality” is conceptually dependent on “true”, and not the other way round) (Habermas, 1973/1989, p. 133, cf. 2002, p. 87; Habermas & Luhmann, 1971/1979, p. 124), a *discourse*-theory of truth would be free from such a disadvantage. A proposition’s *being true* would thus not be elucidated in terms of its correspondence with reality, but in terms of its being agreed upon in the context of a rational argumentative exchange (a version, *mutatis mutandis*, of C. S. Peirce’s epistemic notion of truth). In Habermas’ own words:

If every other person who *could* enter into a dialogue with me *would* ascribe the same predicate to the same object, *then and only then* [*dann und nur dann*] I may ascribe a predicate to an object. In order to distinguish true from false statements, I make reference to the judgment of others—in fact to the judgments of all others with whom I could ever hold a dialogue (among whom I counterfactually include all the dialogue partners I could find if my life history were coextensive to the history of mankind). The condition of the truth of statements is the potential agreement of all others (Habermas, 1973/1989, pp. 136-137, cf. also 2002, p. 89; Habermas & Luhmann, 1971/1979, p. 124).

Let us assess how this first option fares against Rorty’s counterexamples.

2.4. *Returning to Rorty's counterexamples*

We must inquire whether the definition of truth as it appears in these early Habermasian texts provides a basis to the deontological claim that, as truth-seekers, we must be interested in defending our beliefs against everybody – without any limitations of the community of justification in terms of the *competence* of our interlocutors. In fact, what comes immediately to the fore is that the kind of connection between truth and universality that might emerge from such texts had already been questioned, before Rorty, by Albrecht Wellmer in his 1986 *Ethik und Dialog*.

Although the purpose of Wellmer's remarks was specifically to discuss Habermas' discourse *ethics* (as it is developed in Habermas, 1983/1990), and, as a consequence, he focuses on whether or not Habermas' norms for arguments can be considered as having a "moral content" (Wellmer, 1986, p. 105), what matters to our purposes here is the step by means of which he discards this possibility. Wellmer believes that the reason why these norms are not moral ones is that they cannot be understood as norms concerning when to begin or end arguments, and this, in turn, follows from the fact that our obligation, as rational subjects, not to *suppress* arguments, not to ignore objections, *does not entail* that we are obliged to take part in argumentative exchanges with everyone (Wellmer, 1986, p. 105). In other words, if Habermas is proposing a deontology according to which, as truth-seekers, we have such an obligation to discuss *with everyone*, the basis for it is very weak.

According to Wellmer, it would only be irrational to refuse to enter in such exchanges if this were the sign of some kind of *defensive* reaction, as a result of our "fear" of the arguments of others (Wellmer, 1986, p. 106). But, of course, such fear need not be the only reason to refuse entering in certain argumentative exchanges: as Mattias Iser commented,

The obligation to engage in a discourse with anybody who demands it seems extremely counter-intuitive. Just imagine a person on the street who addresses you with the following words: "Dear Sir, the colour of your jacket does not fit that of your pants. I propose that you take the jacket off" Neither would we feel obliged to follow this proposal nor would we be inclined to engage in a discourse with this person (Iser, 1998, p. 153).

Iser's point can be generalized: it is not only that we can discard certain specific propositions as not worth discussing; we can also – and

this is precisely Rorty's point – rationally assume that certain subjects are unable to offer us any interesting arguments. It is that assumption, and not “fear”, which discourages us from discussing trigonometry with children or morality with Nazis. Our responsibility as truth-seekers seems to be *not to protect* our beliefs; we must *face the risk* of our beliefs being refuted – but this does not mean that “anybody” is a bearer of potential refutations. Rorty's point is that there is a wide range of people that we consider too ignorant or too wrong to raise any interesting points about our beliefs. If, as truth-seekers, we had an obligation to discuss moral beliefs even with Nazis and trigonometry with three-year-olds, this should be the outcome of having reasons to believe that interesting objections may come from them. But why should we believe *that*? In other words, the epistemic obligation not to protect our beliefs from possible refutations has to be complemented with empirical considerations about where those possible refutations might appear. And there is no evidence that every human being, simply because of membership in our species, is a possible source of objections to each belief we have. Specifying that our only obligation is to discuss with *competent* interlocutors seems, on the contrary, a reasonable way to describe our epistemic duties.

It is important to notice that the reason to describe them in this way is not *that our time is limited*, that, for instance, children are “low priority” interlocutors. If this were the case, the objection could again be raised that temporal limitations are only part of the empirical circumstances in which our duties are to be carried out, and this empirical point does not affect the obligations we “in principle” have. If the question were one about availability of time, then we might still be obliged to discuss morality with Nazis and trigonometry with children at those occasions in which nothing more pressing requires our attention. However, the problem is not at all one of time. Instead, it is simply that expecting valuable objections about trigonometry to come from three-year-olds, or about moral beliefs from Nazis, may be as reasonable as expecting them from animals or inanimate entities. In other words, *either* conditions such as knowledge and intelligence are somehow irrelevant – and, in this case, there is no reason to stop at the boundaries of our biological species in our quest for potential objections – *or* it is precisely because of fulfilling such conditions that human beings count as potential bearers of objections – and in this case, of course, a reference to *competence* is needed³.

³ This kind of considerations is what prevents us from accepting Daniel Kalpokas's defense of Habermas' position against Rorty's criticism. Kalpokas takes aim on Rorty's

2.5. *The 1996 turn: from a theory of truth to a theory of justification*

Now, this “first round” of the debate concerning the relationship between Habermas’ universalism and Rorty’s counterexamples refers solely to the texts in which the German philosopher attempted to *define* “truth” – that is, to the texts in which Habermas’ purpose is what we have identified as option (A) above. However, as is known, the attempted definition presented in the texts from the 1970s is rather troublesome, which is why Habermas moved to what we have distinguished as option (B). Let us analyze this alternative as well.

Whereas objections such as those raised by Rorty concern the alleged *necessity* of certain consensuses for a proposition *p*’s possibly being true (that is, why should it matter whether or not we agree, even in “ideal conditions”, with *the Nazis*?), insofar as Habermas’s texts also refer to *sufficient* conditions, they have been subjected to symmetric objections: why should consensus under ideal conditions be considered *sufficient* for truth, why cannot there be a rational consensus about, nevertheless, a *false* proposition? This kind of objection was raised, for example, in the context of Habermas’ interview with *New Left Review*: doesn’t the criterion of *Theorie der Gesellschaft oder Sozialtechnologie* entail

tenet that the charge of “performative contradiction” cannot be employed against an exclusivist speaker who wants to “offer *arguments* to the effect that” certain people, such as “feminists, atheists, homosexuals, blacks, etc.” cease to be “taken seriously as moral equals and conversational partners” (Rorty, 2000, p. 8). Whereas the American neo-pragmatist denies that there is any inconsistency at play in such an exclusivist stance (because the exclusivist can always justify his position by saying such things as “women have a distorted view of reality, and the like” (Rorty, 2000, p. 8)), Kalpokas doubts, on the contrary, that “a bigot can argue with those who are to be left outside his community, while at the same time denying his communicative intention, his attempt to convince by means of arguments, and so on” (Kalpokas, 2005, p. 135). We can agree with Kalpokas that such a bigot would indeed be committing a performative contradiction, in arguing *with*, and not only *about*, people that the bigot is, by means of the propositional content of his utterances, declaring unworthy of participation in his community of justification. But this scenario misses the point: what matters here is not what the bigot says but *who he argues with*; the bigot can avoid the performative inconsistency that Kalpokas reproaches, simply by not arguing *with*, say, women, Jews, or homosexuals, but simply arguing *about* them, with interlocutors who are themselves male, Aryan, or heterosexuals. Of course, Rorty would agree that *this* kind of exclusionary attitude is repugnant, but he could also insist, once again, that we reject *this* kind of exclusion precisely because we accept women, Jews, and homosexuals as potential conversational peers, people from whom we can expect reasonable objections. If we do not attribute to some individuals, to begin with, the status of being such peers (as we all do regarding *children*), then we will not need to argue *with* them, but only *about them*, and of course no performative contradiction will emerge from this.

the implausible consequence that an ideally rational argumentative exchange between ancient Greeks would have reached the truth about, say, physics? (Habermas, 1985b, p. 86). Similarly, it is this sufficiency condition which is questioned by Wellmer in *Ethics and Dialogue*: why should we believe that a rational consensus between physicists of the 19th century about Newtonian physics was a proof of its truth—so that, consequently, we can only say that Newtonian physics is not true if we say that there was never a *rational* consensus about it? (Wellmer, 1986, p. 71). If Habermas' way out consists, predictably, in insisting on the fact that truth amounts to rational consensus *under ideal conditions*—which can include the notion that enough time has passed, that is, that subjects attaining this consensus find themselves “at the end of the inquiry” whereas 19th century physicists did not—then the obvious problem emerges that such notion of a *final* ideal consensus tells us nothing useful, in terms of a criterion, about the agreements actually arrived at in argumentative exchanges in the real world. Which finally leads Habermas to accept precisely the kind of dilemma that Wellmer had presented. In Habermas' words,

Either the normative content of the pragmatic presuppositions of rational discourse is insufficient to rule out the fallibility of a consensus discursively attained under approximately ideal conditions. Or the ideal conditions of rational assertibility that are sufficient for this lose the power of regulative ideals to guide behavior because they cannot even approximately be met by subjects capable of speech and action as we know them (Habermas, 2014, p. 38).

Therefore, Habermas no longer believes that “truth” is an epistemic concept, that there is a conceptual connection between truth and justified assertibility under ideal conditions such that the latter notion can be used to *define* “truth”. In the “Introduction” to *Truth and Justification*, he writes:

To be sure, for us there is an unavoidable epistemological connection between truth and justification [...]. But I have [...] become convinced [...] that this does not amount to a conceptual connection between truth and rational assertibility under ideal conditions. [...] Even the arguments that here and now irresistibly convince us of the truth of *p* can turn out to be false in a different epistemic context. [...] These objections have prompted me to revise the discursive conception of rational acceptability by *relating* it to a pragmatically conceived,

non-epistemic concept of truth, but without thereby assimilating “truth” to “ideal assertibility” (Habermas, 2014, pp. 37-38).

Habermas now distinguishes more clearly between, *on the one hand*, the satisfaction of the truth-conditions of a proposition *p* (which satisfaction is not itself an epistemic fact) and, *on the other hand*, what we as epistemic subjects can do to *establish* whether or not such conditions are satisfied. Crucial to this proposal, as follows from the passage we have just quoted, is the circumstance that “establishing” the truth of *p*, or rationally “convincing ourselves” of the truth of *p* does not mean that *p* has been *infallibly* proven true; therefore, we can describe the shift from option A to option B as that from a theory of *truth* to a theory of *justification*⁴.

2.6. The persistence of the problem of competence after the 1996 turn

In any case, what matters to our problem here is that, if not truth, at least the (fallible) *establishment* of truth is elucidated in connection with some notion of universal consensus—that is, the *universalistic* aspect of the problem remains. This comes to the fore in “Richard Rorty’s pragmatic turn”, an article which initially appeared in 1996 and worth quoting *in extenso*:

In [the] 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 defensible 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. This provides the inspiration for the discourse

⁴ In fact, according to Barbara Fultner, what Habermas offered in his texts from the 1970s as a “consensus theory of truth” was *already* “not so much a theory of *truth* as of *justification*. As an account of truth, it is unsatisfactory because warrants can be lost in light of new arguments” (Fultner, 1996, p. 233). She goes on to write that “truth cannot simply be defined as warranted assertibility or rational acceptability since the warrant or justification may be lost, whereas truth is supposed to be an eternal property of a statement” (Fultner, 1996, p. 238). However, there is no real disagreement between Fultner’s position in this point and the reconstruction we are offering here: she acknowledges that Habermas in the 1970s *intended* to provide a theory of truth, but she criticizes it as unsatisfactory and suggests that, as a theory of *justification*, it would fare better. Habermas’ abandonment of an epistemic conception of truth can then be described, following Fultner, as his admission of what form his theory should, from the beginning, have taken.

theory of truth: a proposition is true if it withstands all attempts to refute it under the demanding conditions of rational discourse.

However, this does not mean that it is also true *for this reason*. A truth claim raised for '*p*' says that the truth conditions for '*p*' are satisfied. We have no other way of ascertaining whether or not this is the case except by way of argumentation, for direct access to uninterpreted truth conditions is denied to us. But the fact that the truth conditions are satisfied does not itself become an epistemic fact just because we can only *establish* whether these conditions are satisfied by way of discursive vindication of the truth claim – whereby we have already had to interpret the truth conditions in light of the relevant sorts of reasons for the claim in question (Habermas, 1996, pp. 733-734; 2000, p. 46).

As Levine summarizes this move, “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” (Levine, 2011, p. 246; cf. also Lafont, 2018, p. 301). Therefore, *on the one hand*, Habermas weakens his proposal by making explicit that the epistemic connection between truth and justification “must not be turned into a *conceptually inseparable* connection in the form of an epistemic concept of truth” (Habermas, 2014, p. 38); *on the other hand*, discursive vindications of a truth claim for *p* are the only way in which we can *establish* whether or not the truth-conditions of *p* actually obtain. As a consequence, even in this weaker form of the theory, the analysis of rational discourse, its conditions and the participants in it remains central in such a way that Rorty’s worries remain applicable. The discursive “resistance” of *p* against attempts to refute it no longer means for Habermas that *p* is true, but it *does* mean that we are justified in claiming that *p* is true; therefore, what kind of consensus, the consensus *with whom*, should be taken to justify us in making such a claim?

Therefore, *both* in an understanding of Habermas’s proposal under option A (as providing a theory of *truth*) and in an understanding of it under option B (as providing a theory of *justification*) the problem emerges of why Habermas does not restrict the community of relevant interlocutors to *competent* ones—which is why Rorty’s worries “cut through” the difference between these two versions of the theory.

Now, as we anticipated, in 1971 Habermas *did* include a competence condition in his discourse theory of truth. Let us now consider the texts from that year and the difficulties they raise.

3. The status of “competence in judgment” in Habermas’s 1971 Gauss Lectures: the competence condition and the problem of circularity

Habermas’ point in the early 70s was to admit that we can indeed define truth by reference to consensus obtained by competent subjects. However, Habermas (unlike Rorty) *also* pointed out that in doing so we face a serious problem. Namely, that (in the opposite direction of the definitional work) the very status of “competent” can be reasonably attributed to a subject only on the basis of a rational discourse on the question (Habermas, 2002, p. 96; the passage corresponds to Habermas & Luhmann, 1971/1979, pp. 134-135). In the context of a dialogue with a critic who demands that a competence condition be included in the theory (as did Rorty, Wellmer and McCarthy), Habermas’ declaration that competence can only be rightly attributed as the outcome of a discourse places the German philosopher in a curious situation: *on the one hand*, his theory has to face a circle between “rational agreement” and “competent subject” when it comes to the concrete application of these concepts to particular cases. *On the other hand*, according to this early version of Habermas’ theory, facing this circle is something that an author such as Rorty *should* do; any other attribution of “competence” seems unjustified to Habermas.

Let us reconstruct how this problem appears in Habermas’ 1971 Gauss Lectures (the text of which, in turn, overlaps in a series of relevant aspects with the formulations in *Theorie der Gesellschaft*). Given that these passages have been surprisingly overlooked in the analysis of the Habermas-Rorty debate, we will need to cite some of them *in extenso*.

3.1. Competence and circularity

In the last of his Gauss conferences, Habermas presents an attempted definition of truth in the sense that we saw in section 2.2, above. However, some pages after having introduced this alleged conceptual connection between truth and the assent of *every* other speaker, Habermas adds the following clarification:

According to the consensus theory of truth, the condition for redeeming the truth of propositions is the potential assent of *all* other persons. Now, as a matter of fact, there are always only a few persons against whose assent I can check my assertion’s claim to validity. The actual assent that I can possibly obtain from a few others is more likely to

be endorsed by further judges, the less we and others see any reason to doubt their competence to judge. Therefore we shall restrict the truth condition that has been introduced counterfactually as follows: I may assert *p* if every other *competent* judge would agree with me in this assertion (Habermas, 2002, p. 95; Habermas & Luhmann, 1971/1979, pp. 124-125).

This last sentence is one that Rorty could have gladly agreed to—in fact, it is hardly distinguishable from what Rorty, as we saw, describes as the “ideal presupposed by discourse”. The divergence, however, appears immediately later, when Habermas asks the question “But what can competence in judgment mean in this context?” (cf. Brand, 1976, p. 287 for an early problematization on this point; Habermas, 2002, p. 95; Habermas & Luhmann, 1971/1979, p. 125). Habermas refers here to Kamlah and Lorenzen, who claimed “that competent judges must be capable of performing appropriate verification procedures, [...] must have expert knowledge” (Habermas, 2002, p. 95; Habermas & Luhmann, 1971/1979, p. 125). Nevertheless, the problem for Habermas is how to determine “what sort of verification procedure is to count as appropriate in a given case”, and “who may claim to be an expert”. “These questions, too”, the German philosopher goes on to say,

must be subject to discourse, the outcome of which in turn depends on a consensus among the participants. Expertise is no doubt a condition that must be satisfied by a competent judge. But we cannot specify any independent criteria for what counts as “expertise”; deciding on the choice of these criteria itself depends on the outcome of the discourse (Habermas, 2002, p. 95; Habermas & Luhmann, 1971/1979, p. 125).

The problem of circularity, according to Habermas, emerges not only for the assessment of claims to *truth*, but also for other validity claims:

The consensus theory of truth makes us aware that it is not possible to decide on the truth of propositions without reference to the competence of possible judges. [...] The idea of true consensus requires that the participants in discourse be able to distinguish reliably between reality and appearance, essence and accident, and is and ought; for only then can they be competent to judge the truth of propositions, the veracity of utterances, and the legitimacy of actions.

Yet in none of these three dimensions can we specify a criterion that would allow for an independent assessment of the competence of possible judges or participants in deliberation. *Rather, it seems as though the competence to judge itself must be judged on the basis of the same kind of consensus for whose evaluation criteria are to be found* (Habermas, 2002, p. 96; Habermas & Luhmann, 1971/1979, pp. 134-135. Emphasis ours).

So, summing up, we seem to need to appeal to the subjects' "competence in judgment" in order to be able to trust that some "actual assent", a consensus achieved in some particular discursive exchange, is rational and thus actually probative of anything at all, but it is also only as a result of rational assent that we can ever say, in a justified manner, that certain subjects are actually competent.⁵

3.2. *Why the circularity problem does not emerge for Rorty*

Of course, the circle would not emerge if we could show that (a) the cognitive competence of certain subjects does *not* depend on such a competence having been proven by means of discursive consensus. Or, alternatively, if we held that (b) we can take some discursive consensuses to justify us in claiming truth for certain propositions without it being necessary to problematize, at the same time, the competence of the subjects involved in such consensuses (for example, by taking some discursive consensuses as probative on the basis of their *procedural* properties).

We can illustrate this opposition by returning to Rorty's example of a surgeon facing a Chinese traditional doctor. To make a clearer comparison, let us speak about a *group* of doctors. According to Habermas' proposal, it would be troublesome to say that practitioners of Western medicine are "competent" subjects for the assessment of a series

⁵ We might also speak here of a risk of *infinite regress*, and not only one of circularity. According to Fultner, "Treating rational consensus as a criterion or condition for truth presents the threat of an infinite regress if not of circularity. For it seems that whether a consensus is indeed rational, whether it guarantees truth, in turn depends on yet another consensus" (Fultner, 1996, p. 240). However, Fultner is considering only the relationship between a first-order question, such as "Is *p* true?", and a second-order question, such as "Was the consensus about the truth of *p* a rational one?" (which would also need to be answered discursively) without referring to the key problem of *what subjects* should take part in each argumentative exchange. In other words, Fultner is correct to point a problem out, but the situation for Habermas' theory is even worse.

of propositions referred to human health, whereas Chinese traditional doctors are not: in fact, the “competence to judge” that those doctors have must itself “be judged on the basis of consensus” and, in turn, we will obtain different consensuses depending on how the community that judges about the competence of these doctors is composed.

Now, it seems rather clear that Rorty leans towards solution (a). Though he does not explicitly address the circle Habermas finds between competence and consensus, he does, nonetheless, take certain beliefs as determining “where we are”, as a starting point that provides us with normative standards for the consideration of other beliefs and, indirectly, for the consideration of the competence of other subjects—which would authorize us, in turn, to call the practitioners of Western science, and not Chinese traditional doctors, “competent”. In Rorty’s words, we “should accept the fact that we have to start from where we are, and that this means that there are lots of views which we simply cannot take seriously” (Rorty, 1984b, p. 11).

Insofar as, according to Rorty, “human beings are centerless networks of beliefs and desires and [...] their vocabularies and opinions are determined by historical circumstance”, the possibility may appear that there is not “enough overlap between two such networks”. When that happens, Rorty continues, “we” do not question “our” beliefs; instead, we can call our interlocutors “crazy”, because our attempts to converse “have made us realize that we are not going to get anywhere” (Rorty, 1991b, p. 191). And this is because—as Rorty wrote in the context of rejecting the notion of “transcultural validity”—“The interesting question is not whether a claim can be ‘rationally defended’ but whether it can be made to cohere with a sufficient number of *our* beliefs and desires” (Rorty, 1990, p. 640, cf. also 1991a, p. 50 for an explicit connection between the notion of ethnocentrism and the criterial value attributed to “our own present beliefs”). This amounts to saying that, in the framework of Rorty’s ethnocentrism, unlike that of Habermas’ proposal, the attribution (or, more clearly, the *refusal*) of the condition of “competent subject” takes place *independently* of the result of any previous argumentative consensus *about* competence. According to Rorty, “we” divide subjects as competent or incompetent as the result of the existence or not of a sufficient overlap with the beliefs “we” already have—and *these* beliefs *are not*, at the moment, problematized.

Now, the question whether or not this “Rortyan” solution is acceptable (that is to say, whether or not we can speak of competent subjects without needing to appeal to a *discursive exchange* in the context of which such a competence would have been proven) depends on how

reasonable *Habermas' own* way out of the circle turns out to be. As we will try to show in the next section, Habermas himself seems to move to solution (b)—that is, he wishes to describe some discursive exchanges as probative of truth by virtue only of their procedural properties, without worrying about the competence of the subjects involved in them— but this solution is far from unobjectionable.

4. The appeal to the notion of “ideal speech situation”

4.1. “Expert verification” and “ideal speech situation” in Habermas’ texts from 1971

Let us then return to Habermas’ *ipsissima uerba* in his texts from the early 70s. We arrived at a seemingly unsurpassable circle. It is of no help, adds Habermas, to think that the circle “could be broken [...] by an ontological theory of truth”, as the correspondence theory, because “none of these copy or correspondence theories has [...] held up under scrutiny”. However, he goes on to say, we cannot simply stop at a pessimistic finding that the circle is irresolvable: “Were this the case [...], it would be hard to understand why we nonetheless assume in every conversation that we can reach a mutual understanding. In fact we are always confident that we know how to tell a rational consensus from an illusory one” (Habermas, 2002, p. 96; Habermas & Luhmann, 1971/1979, p. 135).

By stating this, Habermas has not yet introduced what his own solution will consist of, but this turn of the analysis is already interesting. Habermas’ proposal here is (at least partly) descriptive, and not (purely) normative, in the sense that it does not attempt to assess *whether or not we are justified*, after all, in assuming that we can distinguish “a rational consensus from an illusory one”—an attempt that might consequently reach the skeptical result that we are not in fact justified. On the contrary, Habermas is considering here a *factum*, the *factum* that we always make that assumption, and wants to *explain* it, not to seek for its epistemological foundations. If we lacked the confidence that we can distinguish a true consensus, Habermas continues, “we could not tacitly presuppose the sense of speech that is always already accepted at the metacommunicative level and without which ordinary language communication would be meaningless—namely, its rational character. *This phenomenon requires explanation*” (Habermas, 2002, p. 96. Emphasis ours. Cf. the similar lines in Habermas & Luhmann, 1971/1979, pp. 135-136).

So, predictably, Habermas is no skeptic, and he wishes only to *explain* our communicative practices, not to doubt their rationality. What is at stake is, consequently, whether Habermas' solution constitutes a better explanation of such practices than Rorty's own proposal. Now, what Habermas introduces at this point is the well-known, and very polemical, notion of "ideal speech situation". Immediately after presenting a demand for explanation of the "phenomenon" of always counting with our own ability to distinguish rational agreements, he writes:

I would argue that what explains it is that the participants in argumentation mutually *presuppose* something like an ideal speech situation. The defining feature of the ideal speech situation is that any consensus attainable under its conditions can count *per se* as a rational consensus. My thesis is that only the *anticipation* of an ideal speech situation warrants attaching to any consensus that is in fact attained the claim that it is a rational consensus (Habermas, 2002, p. 97; cf. Habermas & Luhmann, 1971/1979, p. 136).

So we see that, according to Habermas, insofar as we want to conceive our communicative practices as rational, we need to assume that we can in fact distinguish instances of truly rational consensus, even if that assumption has to be a presupposition. Now, what is more important for our purposes here is to highlight the fact that, by means of the appeal to the presupposition of an ideal speech situation, two different problems might be solved, but the way in which Habermas develops the notion of such a situation (by referring only to the *procedural* aspect of the argumentative exchange) can satisfactorily tackle only one of them. *One* of those problems is whether or not certain factually obtained consensus about a proposition *p* has been in fact a *rational* one—in opposition to agreements which stem from coercion or some kind of asymmetry between participants, detrimental for the free play of arguments. *Another* problem, the really interesting one, is whether or not a certain consensus about *p* is in fact a sufficient justification for claiming that *p* is true.

The first problem might well be tackled by analyzing only the formal properties of the argumentative exchange, irrespective of any considerations on the cognitive competence of the participants in the dialogue, but it cannot be all what Habermas is concerned about here. "The idea of true consensus", as we have seen in the texts from 1971, "requires that the participants in discourse be able to distinguish reliably

between reality and appearance [...] for only then can they be competent to judge the truth of propositions” (Habermas, 2002, p. 96; Habermas & Luhmann, 1971/1979, pp. 134-135). In other words, participants in a rational discourse who are *not* able to “distinguish between reality and appearance” would not be able to judge the truth of propositions *even if*, from a procedural point of view, their communicative exchange were flawless. Therefore, if we want to know whether an argumentatively agreement about *p* justifies us in claiming truth for *p*, it is not enough to make presuppositions only about *the procedure* that has been followed; it would also be necessary to make presuppositions about, as McCarthy wrote, “empirical conditions”, such as those “referring to the intelligence, competence, psychological normality, etc. of the participants” (T. A. McCarthy, 1973, p. 150).

Now, in Habermas’ words, a speech situation is “ideal if communication is impeded neither by external contingent forces nor [...] by constraints arising from the structure of communication itself. The ideal speech situation excludes systematic distortion of communication. *Only then is the sole prevailing force the characteristic unforced force of the better argument, which allows assertions to be methodically verified in an expert manner*” (Habermas, 2002, pp. 97-98. Emphasis ours). But what follows from this is that if a speech situation is “ideal”, and consequently participants in it guide themselves by the “unforced force of the better argument”, communication will not be “impeded”; in other words, *if* participants in it are capable of “methodical verification in an expert manner”, as Habermas writes, *then* their capability for such a verification will find no obstacles in communication. But it is one thing to say that, in absence of structural obstacles, this ability for expert verification will find its way, and it is quite another to say that purely structural conditions can work a *creatio ex nihilo* of good arguments. For the capacity of expert verification to function “unimpeded”, it needs to be present in the speech situation in the first place—and this demand is not included among the conditions of an ideal speech situation. Habermas only describes the condition that “all participants in dialogue” must have (1) “the same opportunity to [...] initiate communication and to continue it through speaking and responding or asking questions and giving answers”; (2) the same opportunities “to put forth interpretations, assertions, explanations, and justifications and to establish or refute their claims to validity” (Habermas, 2002, p. 98); they must also have, *as actors* (that is, when they are not participants in *discourse*) (3) the same opportunities to use “representative” speech acts, so that they “can translate their non-verbal expressions into linguistic utterances” and

(4) equally symmetric opportunities “to command and resist, to allow and forbid [...], to answer for one’s actions and demand others to do so”, which “ensures the possibility of withdrawing at any time from contexts of interaction and entering into discourses that thematize claims to validity” (Habermas, 2002, p. 99).

Now, it is quite ironic that, whereas Habermas was well aware of the risk of circularity raised by the notion of “competent subject”, as we saw in section 4, when he introduces the presupposition of an ideal speech situation as a solution for this circularity he seems to have completely forgotten the problem of competence. If it is legitimate to *presuppose* that an argumentative exchange is taking place under ideal conditions in the sense that the structure of communication is symmetrical and lets the exchange be guided by “the force of the better argument”, it is not clear why it would not be exactly as legitimate for speakers to *presuppose* (without prior proof, hence without generating any vicious circle) that their interlocutors in the exchange in question are *competent*.

4.2. The problem in “Wahrheitstheorien”

But the disappearance of the problem of competence from Habermas’s sight gets even worse: in “Wahrheitstheorien” (written in 1972 and first published the following year), several elements appear which are surprisingly similar to those in the texts from 1971—all the more surprisingly insofar as, in this later text, they are introduced without *any* reference to the problem of competence which was central in the previous works.

In effect, in “Wahrheitstheorien” Habermas (a) insists again on the *factum* that we always assume that we can distinguish between a rational consensus and an illusory one; (b) declares, yet again, that without such an assumption we would have to give up the claim for a “rational character” of communication (Habermas, 1973/1989, p. 180; the passages are parallel to those in 2002, p. 96; Habermas & Luhmann, 1971/1979, p. 135); (c) that, as a consequence of this, we need to consider the ideal speech situation as an inevitable presupposition made by participants in argumentation (Habermas, 1973/1989, p. 180; compare with 2002, p. 97; Habermas & Luhmann, 1971/1979, p. 136), and (d) that the ideal speech situation is neither a regulative principle in the Kantian sense nor an “existing concept” in the Hegelian sense (Habermas, 1973/1989, p. 181; see 2002, pp. 102-103; Habermas & Luhmann, 1971/1979, pp. 140-141). But now the context in which all

these elements are introduced is not that of the concern about whether “the participants in discourse can be competent to judge the truth of propositions”, but only the remark that “an argumentatively obtained consensus is a sufficient criterion for the redemption [*Einlösung*] of discursive validity claims if and only if, by means of the formal properties of discourse, the freedom of movement between the different levels of discourse is guaranteed” (Habermas, 1973/1989, pp. 176-177).

This notion of “levels of discourse” must be understood in the context of Habermas’ tenet that, for the “structure of argumentation” to produce “rationally motivated decisions about [...] validity claims”, discourse has to allow a possible “revision” of the entire “language system” in which such claims were initially formulated (Habermas, 1973/1989, p. 174). In Pettit’s words, the argumentative practice “must even tolerate interrogation of the very conceptual scheme within which the original claim was put forward” (Pettit, 1982, p. 214). The “freedom of movement” between such levels of discourse, according to Habermas, is guaranteed by, specifically, the “formal properties” of “an ideal speech situation” (Habermas, 1973/1989, p. 177), and at this point he introduces a fourfold enumeration of conditions of such a situation that is roughly the same which we have seen in previous texts (cf. Habermas, 1973/1989, pp. 177-178). But, *again*, and just as we pointed out in connection with the texts from 1971, it is *one* problem whether a consensus about *p* has been rationally obtained, and *another* problem whether the existence of that consensus somehow justifies us in claiming truth for *p*. It is this second problem that Habermas should be able to solve, insofar as he explicitly and repeatedly declares to be in search of the kind of consensus that would be “a sufficient criterion *for the redemption of discursive validity claims*” (Habermas, 1973/1989, p. 177. Emphasis ours.); that in fact a consensus obtained under the conditions of an ideal speech situation can be considered “a criterion for the redemption of each thematized validity claim” (Habermas, 1973/1989, p. 179); or even more strongly, that “a rational consensus” is “truth-guaranteeing” (he speaks of distinguishing a “*wahrheitsverbürgenden Konsensus*” from a deceiving one) (Habermas, 1973/1989, p. 180). Therefore, the same objection applies as before: *even if* we grant Habermas that an argumentative exchange is only rational if it guarantees “freedom of movement” between “levels of discourse”, this freedom, in absence of a sufficient cognitive competence by participants in discourse will not guarantee an adequate assessment of propositions. Even if no structural constraint prevents participants from thematizing, not only the validity claimed for a specific proposition *p*, but also the broader

“language-system” in which *p* was raised, it might well be the case that such participants are simply unable to make reasonable criticisms at any “level of discourse”. Therefore, the result that “no prejudice remains free from thematization and critique” (Habermas, 1973/1989, p. 177) might not be achieved anyway, and not because of structural constraints to the freedom of participants in discourse, but *because of the cognitive limitations of these participants*. Which is why, summing up, Habermas’ appeal to the notion of an ideal speech situation is insufficient.

5. A remaining obstacle: the problem of disagreement and non-neutrality

5.1. A curious exchange of positions: Rorty and Wellmer, 2010

So far, our analysis about whether or not to include a competence condition when reconstructing what kind of consensus has a justifying value with respect to a truth claim has assumed that it is Rorty who defends the inclusion of such a condition, whereas Habermas, though having considered it in texts from 1971, later came to relinquish it. However, even if our proposed description of the debate is, in acknowledging the role of these early Habermasian texts, more nuanced than the one presented by Rorty himself, we have so far omitted any reference to the circumstance that, in the exchange between Rorty and Wellmer in the volume devoted to the former in the “Library of Living Philosophers” series, Rorty casts some doubt on the usefulness of a reference to competence. Let us briefly dwell on this exchange.

In an early criticism of Habermas’ discourse theory of truth, Wellmer had declared that such a theory is false if it does not specify “that true validity-claims are precisely those upon which an uncoerced consensus can be achieved *among those who are sufficiently capable of judgement*” [*Urteilsfähigen*], but that, at the same time, if Habermas’ theory included a reference to such “capacity for judgement” [*Urteilsfähigkeit*], then it “would be void of any substantial content” (cf. also Ferrara, 1987, p. 49, 1989, p. 308; Wellmer, 1986, p. 72. Emphasis ours). However, Wellmer changed his mind on the subject, because, in his contribution to *The Philosophy of Richard Rorty* he made use of the notion of competence precisely in order to defend the tenet that we can consider some reasons to be “good” in spite of the fact that they do not convince everyone. Taking aim at Rorty’s alleged counterexample of a man fruitlessly attempting to convince the prison guards to let him escape to resume his role as leader of the resistance, which, according

to Rorty, proves that there would be no point in the man later saying, “Still, they were good arguments” (Rorty, 2000, p. 9). Wellmer responds that “obviously it would be wrong to conclude that the prisoner did not have good reasons for trying to escape; his prison guards are just the wrong (*an incompetent*) audience to be persuaded by his arguments. The prisoner certainly has a better audience in mind” (Wellmer, 2010, p. 317. Emphasis ours). Indeed, Wellmer continues, when our arguments turn out not to convince our interlocutors, it may well be the case that “Our position may turn out to be weak”, but the case can *also* be that “the audience may be in the grip of strong prejudices or may just be the wrong (an ‘incompetent’) audience” (Wellmer, 2010, p. 318), and so on. “Making a sincere assertion or justifying a belief” is, in the end, “connected with the understanding that the assertion or belief should be acceptable to *any sufficiently competent or enlightened person*” (Wellmer, 2010, p. 327. Emphasis ours).

As we see, what is required for us to consider that a belief, or a reason for a belief, is acceptable is not consensus *tout court*, but consensus among “competent or enlightened subjects”. Insofar as the cognitive competence of our fellow speakers can be questioned, we can still explain their lack of assent to our own beliefs in a way that does not make us doubt about the latter. As we anticipated, it is Rorty who now criticizes the last passage we have quoted: “given the absence of criteria for sufficiency of competence or enlightenment, to make [that] assertion seems just a way of insulting people who disagree with me” (Rorty, 2010, p. 345).

This reply seems odd. What exactly can Rorty mean when he allegedly verifies the “absence of criteria for sufficiency of competence or enlightenment”? Surely enough, Rorty does have in mind something similar to such “criteria” when he decides that there exist, as we saw, “people whose [...] inability to agree with us raises no doubt in our minds about the correctness of our own beliefs”. *Mutatis mutandis*, these people seem to be, from the point of view of Rorty’s ethnocentrism, in the same position as those who, according to Wellmer, are not “sufficiently competent or enlightened”—that is, people whose lack of assent with our reasons would not make us believe that perhaps we are not mobilizing “good reasons” after all. Here again, it is not clear why Rorty’s rejoinder to Wellmer could be legitimate without, at the same time, jeopardizing central aspects of Rorty’s own ethnocentrism. There is, nonetheless, a sense in which his remark on the “absence of criteria” becomes useful to illuminate a last aspect of the Rorty/Habermas debate.

Namely, what Rorty *can*, in fact, claim is that there is an absence of *neutral* criteria “for sufficiency of competence or enlightenment”—in

other words, there may well exist such criteria, but they are subject to disagreement, to deep cultural differences between communities, as can be illustrated by the passage about Western and Chinese traditional doctors. Rorty could in fact accept a principle such as “we are justified in claiming truth for a proposition p if p has been rationally agreed upon by *competent* subjects”, but he would immediately complete this elucidation by adding that who we consider competent will not be a neutral affair; it will depend on some other epistemic commitments. Let us see what the repercussions of this last element are for the debate with Habermas.

5.2. The problem of non-neutrality as a possible argument against the preservation of the competence condition

With this last aspect, the problem emerges that the character of “subject competent in judgment” will not be attributed to the same individuals by (to take up Rorty’s examples) Christian Scientists, traditional Chinese physicians, and secular Western scientists. There seems to be an inescapable *disagreement* as to who deserves to be considered “competent”. Members of a community of Western secular scientists can call themselves “competent” and exclude those of a community of Christian scientists, and vice versa. Therefore, a consequence of including the competence condition in an elucidation of what it is to have justified a proposition is that our analysis of justification will take a form as “we are justified in claiming truth for p if and only if there has been a rational consensus about p by *what our community takes to be* competent speakers”. This is predictably a consequence that Rorty would wholeheartedly accept: in a number of passages, Rorty insists that there is no such a thing as being justified, *period*, but only being justified for “us” (Rorty, 1993, pp. 449-452, cf., for example, 2000, p. 15). Habermas’ universalism, on the contrary, seems *prima facie* entirely at odds with the possibility of accepting a notion of “subject competent to judge” whose attribution depends on previous commitments with beliefs which are distinctive of “our” community. Under this interpretation, it should come as no surprise that Habermas has relinquished a “competence condition” and turned, instead, to the *procedural* characteristics of a speech situation in order to elucidate what it means to be justified in claiming truth for a proposition. In other words, the inescapable problem of non-neutrality in the attribution of competence would create a dilemma for Habermas, who would have *either* to drop the problem of “competence in judgment” *or*, by accepting

that there is no such thing as a neutral attribution of competence, to adopt “ethnocentrism” in the Rortyan sense.

Now, *if* these two were the only options available, then it is clear that Habermas could not make room for a competence condition without assimilating his own position to Rorty’s. However, it should also be clear that the overall difference between Rorty’s epistemology and Habermas’ cannot depend on acknowledging or not that any decision between beliefs belonging to different belief-systems (such as the beliefs of a Western scientist and those of a Chinese traditional doctor, a decision which in turn determines who we will call “competent”) will be based on standards which are not “neutral” between those different belief-systems. In fact, Habermas gladly accepts, when analyzing the operation of “understanding” carried by a social scientist in contact with a radically different culture, that he or she will *need* to appeal to her own standards *against* those of the community he or she is trying to understand, if the two sets of standards happen to be incompatible. Far from refusing to embrace non-neutrality, Habermas writes that even to *understand* an expression, “the interpreter must bring to mind the reasons with which a speaker would if necessary and under suitable conditions defend its validity”; that, in turn, *describing* these reasons “demands *eo ipso* an evaluation” (Habermas, 1985a, pp. 115-116) of such reasons and, finally, that this evaluation is something that the scientist can only do by appealing to *his or her own standards*: “An interpreter cannot [...] take a position without applying his own standards of judgment, at any rate standards that he has made his own. *These relate critically to other, divergent standards of judgment*” (Habermas, 1985a, p. 116. Emphasis ours). Therefore, if, according to Habermas, the social scientist will not only be allowed but *compelled* to make such claims as “S’s affirmation that *p* is not well founded on reasons, *and my standards for declaring this are my own, and not S’s*”, there is no reason why he could not accept similar claims such as “S is not a competent subject, *and my standards for declaring this are those of my community*”. If non-neutrality is not viewed by Habermas as a fatal obstacle for his theory of understanding—a theory which he has, in fact, defended against his critics by insisting that understanding simply *cannot* be neutral (cf. McCarthy’s criticisms in T. A. McCarthy, 1984, and Habermas’s reply in Habermas, 1984, pp. 238–239)—then it would be odd and *ad hoc* to insist on such a need for neutrality when it comes to assessing “competence in judgment”.

The move that *would* turn Habermas’ position into Rorty’s is not, in fact, the bare recognition that a lot of assessments must be made non-neutrally. What is really distinctive of ethnocentrism is not the

acknowledgment that “we” may need to appeal to “our” standards, but a *further* tenet, according to which those standards are not themselves justifiable; in other words, that whereas there can be a justification of propositions on the basis of certain epistemic standards, there can be no *meta*-justification of the standards themselves (cf. for example Rorty, 1986, p. 526, for a denial of the tenet that ‘our’ culture is ‘more closely linked to the nature of humanity or the demands of rationality’ than other cultures are). What Habermas reproaches to Rorty is not the admission that “our” standards differ from those of other cultures, but the tenet that “standards of rationality are simply not to be distinguished in type from any other standards used in our culture” (Habermas, 1992, p. 136)—that is, not to be distinguished from mere *conventions*. Now, our point here is not to assess whether Habermas’ trust in the possibility of such meta-justification of standards, which distances him from ethnocentrism, is in fact well founded. That inquiry is beyond the scope of this article.⁶ Our point is, more modestly, to show that the absence of *neutral* “criteria for sufficiency of competence or enlightenment”, and the subsequent need to appeal to those of “our” community, does not necessarily amount to a *new, specific* difficulty for Habermas’s theory; it is, so to say, a problem the theory already has, for independent reasons.

6. Conclusions

Let us recapitulate. As we saw in section 2, criticisms of Habermas’ universalism such as those suggested by Rorty are unfair if they attribute the German philosopher some sort of naïve epistemic deontology according to which we have a “pedagogic” obligation to

⁶ Rorty, for example, offers an alternative reading of what Habermas describes as the development of “elements of reason” in Western culture (Rorty, 1984a, p. 35) – that is to say, a reading in which standards of current Western culture do not appear as susceptible of a meta-justification that shows their objective superiority, in terms of rationality, *vis-à-vis* previous standards. An adequate assessment of the problem of meta-justification would take us beyond the frame of the Rorty-Habermas debate and would require us to review, *inter alia*, Paul Boghossian’s reading of what he takes to be Rorty’s relativism (Boghossian, 2007, pp. 59-94), allegedly centered on the tenet that “different communities may operate with different epistemic systems and that there can be no facts by virtue of which one of these systems is any more correct than any of the others” (Boghossian, 2007, p. 93), Markus Seidel’s uptake of Boghossian’s insights in the frame of what he describes as the “problem of norm-circularity” (Seidel, 2014, pp. 137-191), Martin Kusch’s reply to Boghossian’s and Seidel’s criticisms of Rorty (Kusch, 2017a) and, more generally, Kusch’s defense of “no-metajustification” (Kusch, 2017b). This is a task to which we will hopefully return to in a subsequent paper.

defend our views to literally anyone. The kind of epistemic deontology that *can* be extracted from Habermas' discourse theory is, however, that as truth-seekers we are obliged to discuss with everybody (and not only with competent interlocutors), which is by itself troublesome. Habermas' actual purposes (to elucidate the notion of truth, first, and that of justification, later) appeal to a notion of consensus that should remain connected, as we argued, with that of *competent* subjects. Given that, as we saw in section 3, Habermas himself did appeal, in early texts, to a notion of "competence in judgment", we needed to dwell on the question whether he had good reasons to relinquish the competence condition in later works. As we tried to prove in section 4, whereas Habermas himself, in the context of his Gauss Lectures, identified a difficulty in the appeal to the concept of competent subjects—that is, the possible circle between determining that an agreement has probative value and determining that participants in it are competent—, his own proposed solution to the circle, which consists of the *presupposition* of an ideal speech situation, does not justify giving more weight to "formal" or "procedural" aspects of this ideal situation (as is symmetry between participants) above the condition of subjective cognitive competence: symmetrical chances of mobilizing arguments cannot, by themselves, guarantee that the consensus be guided by "expert verification" if this capacity is not already present in the participants in the dialogue. Finally, we highlighted in section 5 (by means of a detour through Rorty's brief criticism of Wellmer's appeal to the condition of "competence or enlightenment"), that the inclusion of a competence condition raises doubts concerning non-neutrality in the assessment of such competence. Nevertheless, as we tried to argue, this should not raise a *special* difficulty for a theory, as is Habermas', according to which social scientists are obliged to appeal to their own standards of rationality *against*, if necessary, those of other communities.

References

- Boghossian, P. A. (2007). *Fear of knowledge: Against relativism and constructivism*. Oxford University Press.
- Brand, A. (1976). Truth and Habermas' paradigm of a critical social science. *Sociologische Gids*, 23(5), 285-295.
- Ferrara, A. (1987). A critique of Habermas's consensus theory of truth. *Philosophy & Social Criticism*, 13(1), 39-67.
- Ferrara, A. (1989). Critical theory and its discontents: On Wellmer's critique of Habermas. *Praxis International*, 9(3), 305-320.

- Fultner, B. (1996). The redemption of truth: Idealization, acceptability and fallibilism in Habermas' theory of meaning. *International Journal of Philosophical Studies*, 4(2), 233-251. <https://doi.org/10.1080/09672559608570833>
- Habermas, J. (1984). Questions and counterquestions. *Praxis International*, 4(3), 229-249.
- Habermas, J. (1985a). *The theory of communicative action, 1: Reason and the rationalization of society* (T. McCarthy, Trans.). Beacon Press.
- Habermas, J. (1985b, May/June). A philosophico-political profile. *New Left Review*, 151.
- Habermas, J. (1989). Wahrheitstheorien. In *Vorstudien und Ergänzungen zur Theorie des kommunikativen Handelns*. Suhrkamp. (Original work published 1973).
- Habermas, J. (1990). Discourse ethics: Notes on a program of philosophical justification. In *Moral consciousness and communicative action* (pp. 43-115). (Original work published 1983).
- Habermas, J. (1992). *Postmetaphysical thinking: Philosophical essays*. The MIT Press.
- Habermas, J. (1996). Rortys pragmatische Wende. *Deutsche Zeitschrift für Philosophie*, 44(5), 715-742. <https://doi.org/10.1524/dzph.1996.44.5.715>
- Habermas, J. (2000). Richard Rorty's pragmatic turn. In R. Brandom (Ed.), *Rorty and his critics* (pp. 31-55). Wiley Online Library.
- Habermas, J. (2002). *On the pragmatics of social interaction: Preliminary studies in the theory of communicative action* (B. Fultner, Trans.). The MIT Press.
- Habermas, J. (2014). *Truth and justification*. John Wiley & Sons.
- Habermas, J., & Luhmann, N. (1979). *Theorie der Gesellschaft oder Sozialtechnologie. Was leistet die Systemforschung?* Suhrkamp. (Original work published 1971).
- Iser, M. (1998). Habermas on Virtue. In *The Paideia Archive: Twentieth World Congress of Philosophy* (pp. 148-154). Philosophy Documentation Center. <https://doi.org/10.5840/wcp20-paideia19986146>
- Kalpokas, D. (2005). *Richard Rorty y la superación pragmatista de la epistemología*. Ediciones del Signo.
- Kim, K.-M. (2011). Habermas on understanding: Virtual participation, dialogue and the universality of truth. *Human studies*, 34(4), 393-406.
- Kim, K.-M. (2014). Beyond justification: Habermas, Rorty and the politics of cultural change. *Theory, Culture & Society*, 31(6), 103-123.
- Kusch, M. (2017a). Epistemic relativism and pluralism. In *Epistemic*

- pluralism* (pp. 203-228). Springer.
- Kusch, M. (2017b). Epistemic relativism, scepticism, pluralism. *Synthese*, 194(12), 4687-4703.
- Lafont, C. (2018). Communicative rationality: Vorbereitende Bemerkungen zu einer Theorie des kommunikativen Handelns (Preparatory remarks for a theory of communicative action, 1971). In C. Lafont, H. Brunkhorst, & R. Kreide (Eds.), *The Habermas Handbook* (pp. 288-305). Columbia University Press. <https://www.jstor.org/stable/10.7312/brun16642>
- Levine, S. (2011). Truth and moral validity: On Habermas' domesticated pragmatism. *Constellations*, 18(2), 244-259.
- McCarthy, T. A. (1973). A theory of communicative competence. *Philosophy of the Social Sciences*, 3(2), 135-156.
- McCarthy, T. A. (1984). Reflections on rationalization in the theory of communicative action. *Praxis International*, 4(2), 177-191.
- Pettit, P. (1982). Habermas on truth and justice. *Royal Institute of Philosophy Lecture Series*, 14, 207-228.
- Rorty, R. (1984a). Habermas and Lyotard on post-modernity. *Praxis International*, 4(1), 32-44.
- Rorty, R. (1984b). Solidarity or objectivity? *Nanzan Review of American Studies: A Journal of Center for American Studies*, Nanzan University, 6, 1-18.
- Rorty, R. (1986). On ethnocentrism: A reply to Clifford Geertz. *Michigan Quarterly Review*, 25(3), 525-534.
- Rorty, R. (1990). Truth and freedom: A reply to Thomas McCarthy. *Critical Inquiry*, 16(3), 633-643. <https://doi.org/10.1086/448550>
- Rorty, R. (Ed.). (1991a). Is natural science a natural kind? In *Objectivity, relativism, and truth: Philosophical papers* (Vol. 1, pp. 46-62). Cambridge University Press. <https://doi.org/10.1017/CBO9781139173643.004>
- Rorty, R. (Ed.). (1991b). The priority of democracy to philosophy. In *Objectivity, relativism, and truth: Philosophical papers* (Vol. 1, pp. 175-196). Cambridge University Press.
- Rorty, R. (1993). Putnam and the relativist menace. *Journal of Philosophy*, 90(9), 443-461.
- Rorty, R. (2000). Universality and truth. In R. Brandom (Ed.), *Rorty and his critics* (p. 30). Blackwell.
- Rorty, R. (2010). Reply to Albrecht Wellmer. In *The philosophy of Richard Rorty*. Open Court.
- Seidel, M. (2014). *Epistemic relativism: A constructive critique*. Palgrave Macmillan UK.

- Wellmer, A. (1986). *Ethik und Dialog: Elemente des moralischen Urteils bei Kant und in der Diskursethik* (1. Aufl.). Suhrkamp.
- Wellmer, A. (2010). Rorty on truth, justification, and experience. In R. Auxier & L. Hahn (Eds.), *The philosophy of Richard Rorty*. Open Court.

Received 26th August 2020; revised 22nd October 2020; accepted 2nd November 2020.