

**Gunnar Skirbekk**

University of Bergen  
Professor Emeritus  
gunnar.skirbekk@uib.no

## THE EPISTEMIC STATUS OF THE PRINCIPLES OF JUSTICE IN HABERMAS AND RAWLS

### Abstract

The debate between Habermas and Rawls that took place in 1990s concerned how philosophy can justify the principles of justice under the conditions of pluralism of different and irreconcilable moral, philosophical, and religious doctrines. The context of the debate was mainly Rawls' *Political Liberalism* and Habermas' *Between Facts and Norms* as well. This paper argues that a wider geo-cultural perspective is pertinent in order to better comprehend the different justification strategies in Habermas and Rawls, concerning the principle of justice. This goes for their different geo-cultural experiences and presuppositions – in short, Rawls living in a self-confident North America in the post-war period versus Habermas' German experience of civilization breakdown. However, it might also be relevant for the assessment of these two strategies in our time, faced with new kinds of geo-political differences and conflicts.

### Keywords:

The Rawls-Habermas debate, principles of justice, validity claims, truth, justification

There has been extensive discussion on the epistemic status of the principles of justice in Habermas and Rawls, ever since the debate between Habermas and Rawls in the 1990s, after Habermas' *Faktizität und Geltung*<sup>1</sup> (1992) and Rawls' *Political Liberalism* (1993). An introduction to this debate is presented by James Gordon Finlayson and Fabian Freyenhagen in 2011. Among later contributions I may refer to Silje Aambø Langvatn (2013), James Gordon Finlayson (2019), and to Krzysztof Kędziora, in *Folia Philosophica*, "Habermas and Rawls on an Epistemic Status of the Principle of Justice" (2019).

---

<sup>1</sup> English translation: *Between Facts and Norms* (1996).



In this paper Krzysztof Kędziora rightly points out that Habermas, in this setting, is eager to discuss Rawls' position, whereas Rawls is more reluctant to discuss Habermas' epistemic claims. Habermas is eager to discuss questions of justification, whereas Rawls is more reluctant to do so.

If so, why? To find out, I think it is useful to take a look at the broader context, to get an overview of what is going on between Habermas and Rawls. In this paper, I shall briefly indicate what I have in mind, and why I see such an overview as relevant and even essential.

### HABERMAS ON VALIDITY CLAIMS, TRUTH, AND JUSTIFICATION

There was a long discussion going on between Jürgen Habermas and his older friend and colleague Karl-Otto Apel, in which Apel, time and again (in various long texts, “mit Habermas, gegen Habermas denken”, cf. Apel, 1998) criticizes Habermas for being unclear and evasive, as to the stronger normative justification for basic principles, namely that performed by self-reflexive transcendental-pragmatic reasoning and argumentation. Apel argues consistently that Habermas remains too vague or even tends to argue empirically rather than philosophically.

To my mind, both positions in this debate – Apel versus Habermas – might be seen as too stiff and dichotomic. Case-oriented, cautious and gradualist “arguments from absurdity”, as in the analytic philosophy of language (e.g. Ryle and Searle), might in my view be more promising, as I have argued in various writings.<sup>2</sup> Moreover, when Habermas started talking about four universal, speech-inherent “validity claims” (*Geldungsansprüche*) – truth, rightness, truthfulness and meaningfulness – I equally find his terms somewhat vague and ambiguous.<sup>3</sup> For instance, what about the question of (relative) conceptual (in-)adequacy? How does it fit in, in this scheme of four validity claims? Moreover, the choice of concepts, of conceptual perspectives, is constitutive for what we grasp by single (factual) statements or by comprehensive theories,<sup>4</sup> and also for normative claims (concerning both values and norms). Furthermore, in discussing fallibilism versus truth and certainty, I would argue that cautious case-analyses are essential (2017, pp. 9–28), both in order to illuminate different cases of “tacit” (act-constitutive) “knowing” and certainty in everyday life (cf. Michael Polanyi and late Wittgenstein), and cases of certified truths in various fields in the

<sup>2</sup> For instance: Skirbekk (1993) and more recently in 2017. As to Habermas' response to Apel, see for instance Böhler et al. (2003, pp. 44–64).

<sup>3</sup> As I have argued various places, e.g., in 2017, pp. 29–46.

<sup>4</sup> Cf. Richard Rorty on “redescription”, and the discussion on “horizons”, from Heidegger and Gadamer to Charles Taylor.

sciences and humanities<sup>5</sup>, in addition to the various cases of “certainty” in terms of “absurdities”, as when “arguments of absurdity” are carefully worked out on different cases – all this, in addition to strict self-reflexive certainties (*Nichthintergebarkeit* and strict *Letztbegründung* in Karl-Otto Apel).

Add to this that Habermas in many ways is a “moving target” – changing his approaches and his themes, e.g. *from* “knowledge constitutive interests” (*erkenntnisleitende Interessen*) to speech-act theory (with universal validity claims), and *further to* theory of law (*Faktizität und Geltung*) where (for one thing) “weak motivation” is compensated by institutionalized legal rules, and even further, driven by the same search for a strengthening of moral motivation (seen as a deficit in discourse theory), *now* looking for moral motivation in “religions” (that is: modern and morally decent religions!), in religions as an institutional practice, based on rituals, in addition to theoretical (universal-pragmatic) argumentation and justification for basic normative principles, and in contrast to philosophical thinking that seemingly remains a cerebral activity (even though philosophy is also a practice, even a socializing, mutually-educative practice – on its best).<sup>6</sup>

Moreover, there are some questions to be raised concerning Habermas’ normative justification (of action rules), thus referring directly to main points in Kędziora’s paper, namely Habermas’ discourse principle (D), here in English (as in Kędziora’s paper): “Just those action norms are valid to which all possibly affected persons could agree as participants in rational discourses”. (This principle is then differentiated into “the moral principle” and “the democracy principle”). In German: “Gültig sind genau die Handlungsnormen, denen alle möglicherweise Betroffenen als Teilnehmer an rationalen Diskursen zustimmen könnten” (Habermas, 1992, p. 138). As Habermas himself immediately adds: these formulations contain ambiguous terms (“Diese Formulierung enthält erläuterungsbedürftige Grundbegriffe”). Here, we should mind one point: Habermas, in German, does not talk about “all possibly affected persons”, but about “alle möglicherweise Betroffenen” – hence, in the German original, he is not referring only to persons, i.e., merely to humans. Literally, he is referring to all those who possibly could be affected, humans and non-humans. So, if we take this wider conception seriously, we do not merely have the question of how to include all those humans who cannot participate in rational discourses: “the hard cases”, in bio-medical terminology” and future generations, but also all those “affected” who are not persons and who are not genetically members of *homo sapiens*. One answer: we could and should have “advocatory representa-

---

<sup>5</sup> Such as the knowledge we do have about human macro-anatomy, which is no longer a field of research, but a field for education in health profession, since we *know it all*.

<sup>6</sup> Habermas on religion in this respect, e.g., in 2005, 2012, 2019.

tions” (“advokatorische Vertretung”, as Apel says), and thus do our very best, in a melioristic sense, for all those who cannot participate themselves, be they humans or not, as long as they are “affected”.

However, it remains notoriously problematic to talk about “possible agreement” for an *unknown* number of humans, as well as for all other animals and living creatures that in various ways could be “affected”. In short, a melioristic and gradualist approach makes sense (trying to include as many as possible, as best we can), with important ecological implications, whereas talking about “all possibly affected” is *conceptually unclear* (for one thing depending on how various living beings are conceived).<sup>7</sup>

In short, there is more to be said about these principles, when discussing the relationship between “justification” and “acceptance” in Habermas. Acceptance, possible agreement; “could agree” (“zustimmen könnten”), ok – but, by whom, and how?

#### RAWLS ON POLITICAL IDEALS AND JUSTIFICATION FROM WITHIN

Briefly stated: yes, Rawls has normative ideals, about free and equal citizens, with a sense for fairness and civil duties, living in well-ordered democratic societies – normative ideals that he does not try to justify philosophically. However, instead, according to Rawls, these normative ideals can be found in contemporary societies, in his time, say, in the US after World War Two, at least to a significant degree, namely, to the extent that instead of starting up, once again, with philosophical discussions about basic principles, a more reasonable and realistic approach is that of a justification “from within”.

Now, there are changes in Rawls’ work, *from* reflections on “the original position” and “the veil of ignorance”, *to* reflections on “overlapping consensus” and “reflective equilibrium”, and *to* his later rewritings and reflections on international justice. Anyhow, let us here stay with his conception of overlapping consensus on basic political issues among reasonable persons with different “comprehensive doctrines” (of a reasonable kind), that is, in a modern society with “reasonable” people and “reasonable disagreement”.

This overlapping consensus goes for basic socio-political structures and also for more concrete cases – as Rawls tries to show, by practicing “reflexive equilibrium” (working out better consistency between one’s conception of concrete cases and one’s basic presuppositions).

---

<sup>7</sup> Assuming that Habermas does not restrict “those affected” to those who physically and mentally and otherwise can participate in rational discussions here and now (excluding all future generations, all “hard cases”, all animals that in various ways can be “affected”).

Rawls does not try to justify any one of these basic normative principles, but restricts himself to justify what he argues to be (or should be) an overlapping consensus that is rational and stable, in the sense that various (reasonable) participants, each of them, find good reasons for the same conclusion, but from more or less diverging normative and conceptual positions (“comprehensive doctrines”). In other words, not a consensus for the same reasons, for identical reasons, but neither a mere compromise, a *modus vivendi* (for the time being) out of conflicting interests. In short, good “public reasons” are reasons that make sense for all (reasonable) participants, independent of their different basic presuppositions.<sup>8</sup> The decisive “public reasons” must in principle be accepted by everyone (who is reasonable).

Rawls tries to justify his theory of justice as fairness among free and equal individuals, in well-ordered societies, not from above, but from within – as he conceives it, in a well-ordered society, with reasonable disagreement on some presuppositions (as in religious matters), but with reasonable (free and equal) citizens.

So, all dependent on how we conceive the notion of “normative justification”, we may argue that Rawls does or does not contribute to a normative justification of his political theory (of a liberal, constitutional, and democratic society, as he sees it).

#### UNDERLYING SOCIO-POLITICAL PRESUPPOSITIONS, AND DIFFERENCES?

With this background, it is tempting to take a step back and raise a few questions about possibly different socio-political experiences and presuppositions for the two of them, for John Rawls, born 1921 in the United States of North America, and for Jürgen Habermas, born 1929, in Germany. In short, intellectually they gained a majority in the years after the Second World War, but in very different settings.

Just an anecdotal reminder: at an international conference in the late 1960s, Richard Rorty rejected any philosophical justification for his liberal democratic ideals, with the statement “I am an American”, whereby Karl-Otto Apel stood up, exclaiming: “but I am a German”. Apparently, for Rawls there was no need, no urge for a universal justification of his political convictions. For Apel, with the experience of a civilizational breakdown, the need for a universally valid justification of basic normative principles was paramount.

---

<sup>8</sup> Similar point, in Habermas (1992, p. 139): “Die ausschlaggebenden Gründe müssen im Prinzip von jedermann akzeptiert werden können”.

Could there be a similarity, for Rawls and Habermas, as to their different existential experiences and their different views on the need for a philosophical justification of basic normative principles? For Rawls, justification from within, that's enough. For Habermas, serious justification by universally valid principles is needed.

Yes, I think so, and I think it is worthwhile recalling such “external” factors, trying to understand what they are at, what is at stake for philosophers like John Rawls and Jürgen Habermas.

If we allow ourselves to make such comments, we may take a step further, recalling how the Cold War ended with the fall of the Berlin Wall and the collapse of the Soviet Union around 1990, fostering a feeling of an “end of history” (Francis Fukuyama, 1989<sup>9</sup>), whereby all peoples were supposed to want to be modern in the Western US-style, as it was conceived at that time, in the very special 1990s – before the new NATO (*North Atlantic Treaty Organization*), with the military intervention in Yugoslavia, with 9/11 and Islamist terrorism, and interventionist wars with devastating consequences in Afghanistan, Iraq, Libya, destabilizing whole regions, triggering mass migration; and, at the same time, with increasing inequalities and protests, also within democratic States. And on the top of it: climate crises and other ecological and institutional challenges, worldwide.

In short, there are reasons to claim that now, in the 21st century, there is a need, an urgent need, to defend and to justify some universal principles for a decent human coexistence in a sustainable world society. In retrospect, the 1990s can be seen as a naively innocent interlude, a time when John Rawls’ “justification from within” made sense – since he was a North American, apparently the aim of all peoples and the “end of history”.

Seen in retrospect, from our present challenges, confronted with peoples and regimes who explicitly reject and detest the modern world with enlightenment and reasonable socio-political institutions, seen in this perspective the strategy of “justifying from within”, whether it is in the United States of North America, or in more or less stable states throughout the world, this strategy has apparently lost much of its credibility.

The Habermasian (and Apelian) approach, looking for universally valid justifications for basic political principles, has gained strength, once again.<sup>10</sup>

<sup>9</sup> Quite another perspective, in Fukuyama, 2014.

<sup>10</sup> However, at the same time, there is in our time also something to be said in defence of “arguments from within”: There are cases of “institution constitutive norms”, norms that are required for a institution to exist. Briefly stated, if people want to live in a modern constitutional democracy and welfare state, based on mutual trust and enlightenment, then most of them have to behave accordingly! Hence, there are “institution-constitutive” principles, in addition to general moral principles, and to contingent legal decisions. In such cases, “arguments from within” are relevant and important – e.g., for the integration of migrants from clan societies who want to settle

However, for Habermas himself, his persistent concern for universal justification, nourished by experiences from the civilizational breakdown during the Nazi regime and the atrocities of the Second World War, seems now to a large extent to focus on a concern for the European Union; moreover, his concern for moral motivation and socio-political challenges in modern secularized societies seems to a large extent to focus on an interest in the (hopefully) positive role of beneficial religious learning-processes throughout history, right into the modern age.<sup>11</sup>

At the same time, he has less to say about challenges like climate crises and politicized religions. Challenges of sustainability and the need for a modernizing critique of religion (in a Kantian sense<sup>12</sup>) are not on the top of his agenda. And reasonably so. Nobody can be expected to “cover it all”. There is a “division of labor” also in philosophy, partly at a generational level.

However, for that very reason, it is important to revisit former discussions, like the one between Rawls and Habermas in the 1990s – to grasp what was at stake, and what is still at stake, but also to see more clearly what we now ought to do next.<sup>13</sup>

## BIBLIOGRAPHY

- Apel, K.-O. (1998). *Auseinandersetzungen*. Frankfurt am Main: Suhrkamp.
- Böhler, D., Kettner, M., and Skirbekk, G. (2003). *Reflexion und Verantwortung*, Frankfurt am Main: Suhrkamp, pp. 44–64.
- Finlayson, J. G. (2019). *The Habermas-Rawls Debate*. New York: Columbia University Press.
- Finlayson, J. G. and Freyenhagen, F. (2011) (eds). *Habermas and Rawls. Disputing the Political*. Edited by James Gordon Finlayson and Fabian Freyenhagen. New York: Routledge.
- Fukuyama, F. (1989). The End of History. *The National Interest*, 16, pp. 3–18.
- Fukuyama, F. (2014). *Political Order and Political Decay*. New York: Farrar, Straus, Giroux.
- Habermas, J. (1992). *Faktizität und Geltung. Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*. Frankfurt am Main: Suhrkamp Verlag.
- Habermas, J. (1996). *Between Facts and Norms*. Cambridge, Massachusetts: MIT Press.
- Habermas, J. (2019). *Auch eine Geschichte der Philosophie*. Berlin: Suhrkamp Verlag.
- Kędziora, K. (2019). Habermas and Rawls on an Epistemic Status of the Principle of Justice. *Acta Universitatis Lodziensis. Folia Philosophica. Ethica–Aesthetica–Practica*, 34, pp. 31–46.

---

in modern constitutional democracies with a general and generous welfare state, where differentiation of roles and cultural modernization are required. Cf. Skirbekk, 2019, pp. 21–22, 35–36.

<sup>11</sup> Cf. Habermas’ *opus magnum*, the impressive two-volume *Auch eine Geschichte der Philosophie* (2019), with more than 1.700 pages, but without addressing a reasonable philosophical critique of religion, as in Peter Rohs 2013 (a Kantian theist), or Holm Tetens 2015 (a rationalist theist), *Gott denken. Ein Versuch über rationale Theologie*.

<sup>12</sup> That is, as purification, not as rejection; recall the titles (in English): *Critique of Pure Reason*, *Critique of Practical Reason*, *Critique of Judgment*.

<sup>13</sup> Cf. the “time-diagnosis”: Skirbekk, 2019.

- Langvatn, S. A. (2013). *The Idea and Ideal of Public Reason. John Rawls' Attempt to Conceptualize a Well-ordered Constitutional Democracy*. Bergen: University of Bergen.
- Rohs, P. (2013). *Der Platz zum Glauben*. Münster: Mentis Verlag.
- Skirbekk, G. (1993). *Rationality and Modernity*. Oslo–Oxford: Scandinavian University Press–Oxford University Press.
- Skirbekk, G. (2017). *Philosophie der Moderne: Vernunft, Wahrheit, Menschenwürde, Meinungsfreiheit*. Weilerswist: Velbrück Wissenschaft.
- Skirbekk, G. (2019). *Epistemic Challenges in a Modern World*. Zürich: LIT Verlag.
- Tetens, H. (2015). *Gott denken. Ein Versuch über rationale Theologie*. Stuttgart: Reclam.