



Steve Fuller*

Symmetry in World-Historic Perspective: Reply to Lynch

<https://doi.org/10.1515/auk-2022-2022>

Abstract: William Lynch has persistently questioned the politics underlying my appeal to science and technology studies' flagship symmetry principle. He believes that it licenses the worst features of the 'post-truth condition'. I respond in two parts, the first facing the future and the second facing the past. In the first part, I argue that the symmetry principle will be crucial in decisions that society will increasingly need to make concerning the inclusion of animals and machines on grounds of sentience, consciousness, intelligence, etc. In the second part, I argue that the symmetry principle has been in fact at the core of the 'justice as fairness' idea that has been at the core of both liberal and socialist democracies. Difficulties start once the means of expression and communication are made widely available and the standards of fairness are subject to continual questioning and renegotiation.

Keywords: Turing, Kant, Rawls, fascism, democracy, symmetry, Neurath, Popper, Weber

Lynch doubts the political soundness of my appeal to a generalized symmetry principle of the sort that has been characteristic of science and technology studies. This principle would adopt a broadly 'neutral' approach to the study and evaluation of any social phenomenon, such that whatever empirical or normative conclusions are reached would not be biased by whether the phenomenon is already deemed 'true' or 'false', 'good' or 'bad', or perhaps even 'normal' or 'strange'. For Lynch, it opens the door to the worst features of the 'post-truth condition' (Lynch 2022). My response proceeds in two parts. The first part updates the normative import of the symmetry principle for the future into which we are heading, where it will become increasingly relevant, even as it acquires a new form. The second part addresses Lynch's historical understanding of the

*Corresponding author: Steve Fuller, Department of Sociology, University of Warwick, Coventry, UK, E-mail: s.w.fuller@warwick.ac.uk

Open Access. © 2022 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

principle, which is closer to the direct concern of his critique of my appeal to the principle.

1 Understanding the Future of Symmetry Now

A good way to think about the continued relevance of the symmetry principle is to consider a recent event. As I write this response, Blake Lemoine, an engineer who works in the Responsible AI division of Google, has been put on administrative leave for having declared that the company's 'Language Model for Dialogue Applications' (aka 'LAMDA') is sentient (Vallance 2022). It's not clear why this drastic decision has been taken, but plenty of 'experts' have been quick to accuse Lemoine of 'anthropomorphism', which suggests that only humans are sentient. But of course, that isn't quite right, since for at least the past 250 years, animals with nervous systems have been also considered sentient. Moreover, the idea that animals are sentient pre-exists knowledge of the nervous system. Much earlier, it had been a staple of Epicurean philosophy, even though Epicurus didn't know what it was about animals that made them sentient. However, he could tell the difference between sentient and non-sentient beings based on their demeanour. The Epicureans speculated about how the complex organization of atoms might produce sentience, which two millennia later eventuated in neuroscience. But before the mid-eighteenth century, judgements of sentience had been based on observations of behaviour. Thus, displays of pain have loomed large as criterial of sentience. In this respect, 'anthropomorphism' simply signals the fact that the arbiters of sentience always seem to be human, since in principle some animals should be able to judge pain in humans as well as humans can judge pain in animals, especially given similarity in the background physiology of the two beings.

Of course, one could argue that even the most sentient non-human animals lack the sort of mental presence, or 'consciousness', that would enable them to pass interspecies judgement. But this is not so obvious in the case of an advanced AI like LAMDA, notwithstanding its lack of a physiology comparable to sentient animals. After all, unlike animals, AI machines are purpose-made to communicate with humans. And because judgements of sentience are based on behavioural displays, what matters is that the entity can communicate in the display medium. So, if LAMDA can express what humans recognize as pain and can recognize what humans express as pain, then why would it not count as sentient? Would we hold against LAMDA the mere fact that it's made of silicon and programmed by an algorithm rather than made of carbon and programmed by a genome? Interestingly, the 'experts' who were quick to condemn Lemoine

raised the precedent of ELIZA, the relatively primitive algorithm that nevertheless passed in the 1960s as an adequate psychiatrist to those who consulted it to relieve their mental distress. For the experts, ELIZA exemplified the unwarranted projection of human qualities onto machines. But of course, they were not the ones seeking psychiatric advice—and those who dealt with ELIZA found the machine's responses as illuminating as those of a human.

Several features of 'symmetry' in the sense that divide Lynch (2022) and me are at play here. First, a standard of judgement always requires construction because the case that demands judgement is rarely under the control of the judge. Rather, a judge (or jury)—often selected randomly—is called upon to decide the relationship between the case and a more general principle, as in 'Does the act for which the defendant has been brought to trial constitute a crime of a certain sort?' The judge then sets the *criterion* by which the answer is to be determined. The highlighted term, a staple in modern epistemology, was introduced by the Stoics as an extrapolation of the Athenian judicial system. The criterion provides the terms of engagement for conducting the trial, which in turn secure the grounds for justifying the outcome of the case. However, as the criterion has passed from the law to philosophy, the judge's task has been seen more neutrally as one of classification—that is, to find a basis for either including or excluding the case vis-à-vis some assigned concept.

Such was the original 'anthropological' appeal to symmetry in the late imperial period (e.g., Edward Evans-Pritchard and his student Mary Douglas), which then anchored anglophone rationality debates of the 1950s and '60s, and which David Bloor (1976) later adapted as the signature methodological move of what became science and technology studies. In this context, it was important for judgement not to be biased by whatever other, typically 'historical' knowledge that the judge might have that would distinguish the case at hand from previously classified cases. The politics of this approach to judgement is reasonably seen as 'postcolonial' *avant la lettre*: It went against the tendency to anchor paradigms of reason in Western practices, as if a culture must have taken a path like the West's if it is to be deemed 'rational'. It is sometimes forgotten nowadays that this ethnocentric presumption was shared by many mid-to-late twentieth century social science theories, ranging from Joseph Needham's claims for the uniquely European character of science to Jean Piaget's brand of cognitive psychology, which made it seem that children 'naturally' recapitulate the history of Western science and morals—that is, unless arrested by their home cultures (Needham 1972).

Bloor's historian followers (e.g., Shapin and Schaffer 1985) manifested this 'postcolonial' concern by refusing to invoke ways of understanding the past to which the past agents themselves would not have had access—especially knowledge of the extent to which their beliefs and actions are now seen as vindicated. It

would amount to the present exerting illegitimate control over our understanding of the past. But of course, the problem could happen in reverse: namely, that the past might exert illegitimate control over our understanding of the present. In the case of LAMDA, this might come from the fact that we have traditionally attributed sentience to beings with nervous systems, something that LAMDA lacks. In that case, a ‘symmetrical’ treatment of LAMDA as a candidate for sentience requires that the criterion be indifferent to how the sentience might be produced—that is, the path by which the candidate might have taken to be deemed sentient. In short, sentience is rendered a second-order concept, or ‘functionalized’ in the sense that Ernst Cassirer (1910/1923) originally identified as the great intellectual leap taken first by Galileo in physics and later Frege in mathematics.

The underlying metaphysical intuition identified by Cassirer is that sentience is not something possessed by individuals as a property (*à la* Aristotle) but rather a style of being in the world that in principle all individuals might display to varying degrees, depending on context. In that case, the task of a ‘criterion’ of sentience—true to the word’s Greek etymology—is to establish a basis for making the cut whereby a candidate is ‘sentient enough’ to pass. This articulation of the symmetry principle should conjure up not only the Turing Test but also the ‘veil of ignorance’ by which Rawls (1971) claimed to derive the principles of a just society. Common to both is a sense of ‘fair play’, which in turn implies setting up the game of ‘passing’ so that, in principle, any eligible player could act as a referee. While this point is often—and rightly—seen as a recasting of Kant’s categorical imperative, the intuition can be given a more future-facing ‘temporally symmetrical’ facelift, whereby the ever expanding ‘cosmopolis’ would allow those who pass Kant’s test for humanity later in history to pass judgement on those who had passed earlier (Fuller 2022). In other words, those who passed as ‘human’ in the past may be judged to have been ‘inhumane’ in the future—as well as vice versa.

2 Looking Back at What Lynch Thinks is Wrong with Symmetry

I am sceptical of Lynch’s portrayal of the normative character of the contemporary political landscape. In particular, I reject the label ‘Fascism’ as appropriate to post-truth populism, mainly because of what people who use the label normally mean by it. They mean disrespect for established institutions to the point of encouraging and committing violence against them, as exemplified by the 6 January 2021 insurrection at the US Capitol. In this context, Donald Trump is portrayed as a dog whistling ringleader, whose ultimate aim was to be reinstated as

US President, presumably with a larger authoritarian—and perhaps even totalitarian—agenda in tow. This hidden agenda is often portrayed as ‘racialized’, given that the core of Trump’s support comes from politically and economically disenfranchised whites, especially males. However, the disenchantment with established institutions predates Trump and crosses racial and ideological divides. Indeed, as Lynch himself acknowledges, the supporters of Trump and left-wing Bernie Sanders have overlapped, but Trump made more headway because he operated within a political party (Republican) with a depleted establishment. In any event, Trump made more visible—if only as a lightning rod—simmering frustration with the ‘long game’ social democratic/neoliberal-style of piecemeal remediation of race- and gender-based injustices. In this regard, Hillary Clinton’s 2016 presidential pitch to continue the ‘progressive’ legacy, specifically building upon Obama’s ‘achievements’, triggered opposition, often spilling over into contempt. And whereas Clinton waged a relatively conventional campaign in terms of how it addressed and studied voters, Trump, armed with Breitbart impresario Steve Bannon, commandeered social media, effectively changing the rules of the game for how presidential elections are fought.

The one—and certainly not trivial—aspect of this story that resembles historic Fascism is the mobilization of emerging media in a population that is increasingly free in terms of both intellectual access and expression. Here it is worth recalling that coinciding with Germany’s perceived humiliation in the First World War settlement was the rise of radio, tabloid newspapers and the film industry. Berlin was at the heart of these developments, much of which was fuelled by foreign investors taking advantage of Germany’s artificially depressed economy. Meanwhile the constitution drafted in Weimar replaced the Kaiser with a republic that aimed to maximize individual and collective representation in public life. Taken together, the situation was bound to be incendiary—and was recognized as such at the time. The 1920s was marked by much discussion about who or what might turn the ambient instability to their advantage. The first of Fritz Lang’s classic series of silent films featuring the diabolical speculator Dr. Mabuse epitomized the spirit of the time. Its logic and aesthetic would receive a transatlantic makeover in the 1930s as the Batman franchise, in which the story is told from the standpoint of the hero rather than the villain.

The point worth underscoring here is that until the rise of Hitler, the sort of person thought to be poised to capitalize on German turmoil was a financier with superhuman powers. In today’s terms, a venture capitalist such as George Soros or Peter Thiel might fit the bill, depending on the beholder’s political orientation. Indeed, until he became Chancellor, Hitler’s critics in the intelligentsia generally took him to be in the pocket of Alfred Hugenberg, the great media mogul of the Weimar Republic who helped to craft Hitler’s image and briefly served as his

Economy Minister. However, it was *against* this stereotype that Hitler decisively asserted his populist credentials by quickly side-lining Hugenberg and putting the capitalist establishment on notice that he was running the show as someone whose legitimacy rests on facing the people directly rather than from behind the shadows.

We should not underestimate the epistemic significance of this move, which links transparency with authenticity as the royal road to truth in a democratic society. To be sure, Lynch wants to replace ‘democratic’ with ‘populist’, and I would agree, if Lynch were treating the two terms as synonyms. But clearly, he sees them somehow as closer to antonyms. Yet, one of Lynch’s own exemplars of good democratic practice from the Weimar period, Otto Neurath, promoted just this link between transparency and authenticity through a spatio-temporally specified ‘observation language’ in the Vienna Circle, whose public face was manifested in ‘Isotype’ pictorial social statistics. A latter-day descendant of this mentality, for better or worse, has been Donald Trump’s endless stream of tweets in real time response to world events. Trump’s spontaneity, which conforms to the practice of the more than 200 million daily Twitter users worldwide, demonstrated that he wasn’t hiding behind the shadows of a cabal of conspirators pulling his strings. Indeed, most of the conspiracy theorizing that has emanated from Trump’s camp is about the establishment-based ‘deep state’ trying to undermine him—and by extension, the will of the American people.

Here we need to think of Twitter’s self-imposed constraints as a social media platform—at least before it ousted Trump—as the technological equivalent of logical positivist ‘protocols’ in a world devoid of a common language of thought, yet at the same time possessed of many locally available ways of securing at least rough translation (e.g., Google Translator). Perhaps Lynch would find such an epistemically luminous portrayal of Twitter inappropriate, if not whimsical. If so, he would be short selling the social-epistemic radicalness of logical positivism. After all, while Neurath may not have stormed the US Capitol, he offered his services to a self-styled ‘Soviet’ breakaway republic in Bavaria at the start of the Weimar period. For this he was put on trial, during which Max Weber attempted to excuse Neurath for being an overenthusiastic technocrat. The result was Neurath’s expulsion from Germany—and his loss of employment as Weber’s assistant in Heidelberg—but just in time to be a founder of the Vienna Circle. In this context, Lynch ignores the similarity between what Robert Proctor dubbed Neurath’s ‘Neutral Marxism’ and the generalized symmetry principle from which he recoils (Proctor 1991, 168).

Here it is worth recalling that Neurath co-translated (with his wife) Francis Galton’s *Hereditary Genius* into German before the start of the First World War and was a fellow at the US-based Carnegie Endowment for International Peace

shortly thereafter, where, along with fellow economist John Maynard Keynes, studied how the gains in productivity, innovation and overall efficiency manifested by states on a war footing could be extended to peacetime prosperity (Fuller 2018, 83). Both projects appeared to be geared towards making a virtue out of a necessary evil. Neurath's 'neutrality', such as it was, involved carefully separating means and ends: Just because a certain means arose in connection with the pursuit of bad ends, it doesn't follow that the same means could not be redeployed for good ends. This clearly dualistic ends-means approach to rationality undoubtedly endeared Neurath to Weber, notwithstanding his disdain for Neurath's socialist politics. Many staples of positivist philosophy of science, such as the strong distinction between the contexts of discovery (aka 'ends') and justification (aka 'means'), relate to this sensibility. It effectively defines objectivity as *redeployability*, whereby the same means can be made to work equally well to bring about radically different ends. I have placed this mindset at the core of the post-truth condition (Fuller 2020a, 5).

Weber and his contemporaries typically understood such redeployability in terms of a Nietzschean transfiguration of values, which amounts to treating any practice as neither good nor bad in itself but eligible to become either depending on the value horizon in which it is contextualized (Joas 2000). Thus, a gun is at once something you should never own and just what you need to remove a persistent foe. The Catholic 'doctrine of double effect', much loved by the Jesuits, was conceived in this spirit and it has proved most useful in the context of justifying war. This position is sometimes cast as radical value subjectivism, if not nihilism. But from the standpoint of post-truth, it's more perspicuous to see it from the other side of the coin, namely, that objectivity lies in the 'rules of the game', which is always about redeploying the resources at play, through which the values of the various players might be realized.

The part of this picture that Neurath probably did not see clearly was that knowing the rules of game doesn't necessarily increase the chances of 'social progress' in the sense that the 'good guys' are better positioned to capitalize on this improved knowledge. Here Neurath's Viennese sparring partner Karl Popper saw more clearly, given his stress on the 'reversibility' of values and outcomes in a just democratic regime, be it political or scientific. If Popper seemed 'precautionary', as Lynch suggests, that was because he wanted to place boundary conditions on the amount of harm that can be done in a world that licenses so much turbulence. Underlying this difference in attitude between Neurath and Popper was the spirit in which the boundary conditions should be set in the first place. There are two ways of thinking about this. (1) While Neurath was fixed on specific foes, against which everyone might be equally willing to mobilize (as in a wartime economy, but one might think of climate change in the same vein today), Popper held

that the foes would always be multiple, and that people would spontaneously organize against them differently. (2) While Neurath thought that a well-ordered society would allow everyone to realize their potential in a way that complemented everyone else's potential (aka Isaiah Berlin's 'positive liberty'), Popper held that such a utopia was a world-historic mirage, and that the best scenario for humanity was a succession of temporary governments of the sort valorised in modern liberal democracies, where one should always expect winners and losers—but not predictably, which is a far from trivial point.

This set of contrasts captures the difference between collectivism and individualism—and what people think they're defending when they talk about 'socialism' and 'capitalism', respectively. I have argued that the clarity of this difference remains obscured due to Marxism's historic muddying of the waters with its passive-aggressive attitude towards capitalism, upon which it aims to build as it undermines the system (Fuller 2020b). Here it is worth recalling that when Weber's great rival Werner Sombart coined 'capitalism' in 1902, he was referring to a society in which the 'rules of the game' do not pre-exist the players—a common feature of both natural law theory and Hobbes' Leviathan—but are made up by the players as they play, primarily through the market's price mechanism. In other words, the difference between the governors and the governed is collapsed in practice as values are permitted to fluctuate in public view (Fuller 2020a, chapt. 4). Of course, Sombart was imagining an idealized capitalist order that does not allow for path dependent forms of success to develop (aka monopolies), but the deeper democratic ideal, whereby anyone could be the judge of anyone else, remains salient—and gets to the normative heart of a generalized symmetry principle that I remain happy to uphold.

References

- Bloor, D. 1976. *Knowledge and Social Imagery*. London: Routledge & Kegan Paul.
- Cassirer, E. 1910/1923. *Substance and Function*. Chicago: Open Court Press.
- Fuller, S. 2018. *Post Truth: Knowledge as a Power Game*. London: Anthem.
- Fuller, S. 2020a. *A Player's Guide to the Post-Truth Condition*. London: Anthem.
- Fuller, S. 2020b. "'Knowledge Socialism Purged of Marx: The Return of Organized Capitalism'." In *Knowledge Socialism. The Rise of Peer Production, Collaboration, Collegiality and Collective Intelligence*, edited by M. Peters, T. Besley, P. Jandric, X. Zhu, 117–34. Berlin: Springer.
- Fuller, S. 2022. "'Kant after Kant: Towards a History of the Human Sciences from a Cosmopolitan Standpoint'." In *The Palgrave Handbook of the History of the Human Sciences*, edited by D. McCallum. London: Palgrave.
- Joas, H. 2000. *The Genesis of Values*. Chicago: University of Chicago Press.

- Lynch, W. 2022. “Does Post-Truth Expand or Restrict Political Choice?” *Analyse & Kritik* 44 (1): 137–59.
- Needham, R. 1972. *Belief, Language and Experience*. Oxford: Blackwell.
- Proctor, R. 1991. *Value-Free Science? Purity and Power in Modern Knowledge*. Cambridge: Harvard University Press.
- Rawls, J. 1971. *A Theory of Justice*. Cambridge: Harvard University Press.
- Shapin, S., and S. Schaffer. 1985. *Leviathan and the Air-Pump*. Princeton: Princeton University Press.
- Vallance, C. 2022. “Google Engineer Says Lamda AI System May Have its Own Feelings’.” In *BBC News (14 June)*.