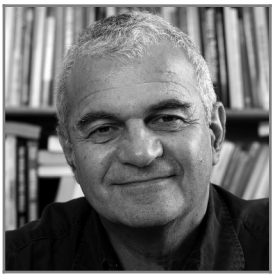
Эпистемология и философия науки
2026. Т. 63. № 3. С. 67–77
УДК 167.7

Epistemology & Philosophy of Science
2026, vol. 63, no. 3, pp. 67–77
DOI: https://doi.org/10.5840/eps202663341

# Reply to critics

**Harry Collins** – FBA.
Honorary Professor.
University College London.
Emeritus.
Cardiff University;
e-mail: CollinsHM@cardiff.ac.uk



I begin my suggesting that the instincts that accompanied early, 1970s, work in the sociology of scientific knowledge – that we were doing something important – have turned out to be sound. Fifty years later responses to an account of those early years lead to deep discussions about artificial intelligence, the nature of sociology and the nature of philosophy. Thereafter, my responses to the discussants are separated into two parts. I disagree with Filatova that recent work in artificial intelligence has produced devices with human-like intelligence and I explain why. She is certainly right in saying that the output of large language models is more readily mistaken for human-like responses than that of previous generations, but this means only that we have to be more ingenious in exploring the ways they differ from us. Thorpe and Sokolova invite a discussion of the trajectories of sociology and philosophy and their relationship to each other. I find that I no longer know whether I am a sociologist or a philosopher though nowadays I have more sympathy with the concerns of philosophy than with what lies at the centre of attention of contemporary sociologists. This invites a discussion of the quintessential sociology and quintessential philosophy, topics which are very unpopular within the constructivist way of thinking. Nevertheless, such an approach is vital if we are to resist the erosion of science and other academic disciplines and, indeed, the erosion of truth itself under contemporary political pressures. I introduce the notion of 'formative aspirations' as a way of defining the essence of science and of activities like sociology and philosophy. I suggest that sociology and philosophy collide in terms of their formative aspirations because both are concerned with the ultimate grounds of truth.

***Keywords:*** Strong Programme, sociology of scientific knowledge, interactional expertise, formative aspirations, philosophy and sociology

# Ответ оппонентам

**Гарри Коллинз** –
член Британской академии.
Почетный профессор.
Университетский колледж Лондона.
Заслуженный профессор.
Университет Кардиффа;
e-mail: CollinsHM@cardiff.ac.uk

Настоящая статья представляет собой ответ Гарри Коллинза на комментарии к работе «Батская школа и Сильная программа». Интуиция, сопутствовавшая ранним работам по социологии научного знания в 1970-е гг., – ощущение значимости предпринятого – оказалась верной: спустя полвека отклики на описание тех первых лет оборачиваются глубокими дискуссиями об искусственном интеллекте, о природе социологии и о природе философии. Ответы участникам обсуждения разделены на две части. Коллинз не согласен с Асей Филатовой в том, что новейшие разработки в области искусственного интеллекта привели к созданию устройств, сравнимых с человеческим разумом, и поясняет почему. Автор признает, что продукты больших языковых моделей ныне гораздо чаще, нежели ранее, можно принять за высказывания человека, однако это лишь означает, что нам следует проявлять больше

© Harry Collins, 2026

изобретательности в поиске того, чем они от нас отличаются. Чарльз Торп и Татьяна Соколова предлагают обсудить траектории развития социологии и философии, а также их взаимоотношения. Коллинз приходит к тому, что более не может сказать определенно, кто он – социолог или философ, хотя сегодня испытывает бóльшую близость к философской проблематике, нежели к тому, что составляет центр внимания современных социологов. Это естественным образом подводит к вопросу о том, что есть подлинная социология и подлинная философия, – вопросу, крайне непопулярному в рамках конструктивистского подхода. И все же подобная постановка вопроса необходима, если мы намерены противостоять эрозии науки и других академических дисциплин, а также – что еще важнее – эрозии самой истины в условиях современных политических реалий. Автор вволит понятие «формативных устремлений» как инструмент для определения сущности науки и таких видов деятельности, как социология и философия. Его гипотеза состоит в том, что социология и философия сталкиваются именно в своих формативных устремлениях, ибо обе обращены к предельным основаниям истины.

***Ключевые слова:*** Сильная программа, социология научного знания, интеракциональная экспертиза, формативные устремления, философия и социология

I thank the three commentators. It is marvellous that a discussion of the early days of science studies could elicit responses that have to do with recent developments in artificial intelligence, with the deep features of sociology, and with the future of philosophy. It felt in those early days that something big was going on and. given the breadth of the responses, it looks like that instinct was sound.

I'll take the responses in the order in which I received them. Filatova sent hers in early and has been, perhaps, unlucky because I disagree with her and she gave me plenty of time to say why. Thorpe and Sokolova invite a different kind of discussion.

## Filatova

> Traditional critiques of AI no longer appear fully convincing, neither to those who design such systems nor to those who engage with them in everyday practice. At the very least, they are not compelling enough to prevent users from interacting with LLMs in contexts where these systems are experienced as meaningful and useful [Filatova, 2026, p. 63].

May I say that Filatova's reply to my piece is not really a *reply* because she talks about artificial intelligence (AI) which I don't mention. But that's OK – I have written about AI since the 1980s (three books and quite a few articles) and I do mention expertise in my piece. What she



says is that recent AI has rendered my analysis of expertise out of date since today's computers must be treated as social actors with their own expertise.

Filatova is unlucky in that the field of AI is moving fast and even since she wrote what she wrote another paper of mine on Large Language Models (LLMs) has come out (in *Synthese*), another has been accepted for publication and a third is under submission but posted on arXiv [Collins. Thorne, 2026; Collins et al., 2026a; Collins et al., 2026b]. These papers all deal with the difference between LLMs and human intelligence showing how LLMs fail or are potentially dangerous. All three of these papers are accessible via the internet. I cannot go into the detailed content here except to say that the 'alignment' with human values that Filatova talks about is as much a danger as a benefit since, as we show, it is easy for malign persons to align these devices in malign ways (see the third paper listed above). Instead, I'll take a more distanced look at our differences in approach.

Going to the quotation at the head of this response, I cannot think of a time when the opposite of what Filatova says would be true. Taking the first sentence, AI is an industry depending on sales to the public and stories for venture capitalists and it is never going to welcome criticism (unlike proper sciences). As for the second sentence, it is true that people do continually find themselves treating AI as expert actors but to claim that shows they have finally achieved human-like intelligence is to look through rose-tinted spectacles. When we look at the philosophy/sociology of knowledge, it doesn't matter what properties humans impute to computers because humans are capable of attributing all kinds of properties to all kinds of things, not only AIs but all kinds of non-humans, such as scallops, as Latour and Callon exemplify [Collins, Yearley, 1992]. Worse, in wartime, groups of humans often attribute non-humanness to other groups of humans to make killing them less stressful. In the field of AI, we can go at least as far back as 1976 and Weizenbaum's Eliza to find humans attributing psychotherapeutic abilities to the simplest of programs. Distinguishing LLMs from humans is harder than ever because of their extraordinary fluency with native language but that's not a philosophical argument, it's an observation of how we normally attribute intelligence to others. Latour, the most thoroughgoing of the attributors, has shown that we can attribute agency to anything, so of course there is going to be a space for exploring how the latest generation of AI affects society – and affect it, it will. But to investigate how computers affect society rather than to investigate their intelligence is Filatova's choice. My view is that the interesting thing for a sociologist of knowledge rather than a sociologist of labour, who might just as easily be looking at the introduction of the spinning jenny, is whether the latest generation have really solved the scientific problem of reproducing human intelligence – something that has been claimed for every new generation of AI from the beginning.

In my 1990s book I spend some time asking whether pocket calculators are, really, better arithmeticians than humans and conclude they are not because they don't know how to approximate. Of course, each new generation of AI is better than the last and LLMs *can* approximate. I use ChatGPT all the time but only in the way I use a pocket calculator which, as I explained in 1990, involves having a good idea of the answer first if you are not to make terrible mistakes. My complaint about Latourianism is that it doesn't allow you to ask the critical questions about non-humans' abilities since it attributes the qualities to non-humans that ordinary humans attribute to them in ordinary social life, so it is impossible to say that anyone is getting anything wrong. That is not a club I wish to join but if that's what you want to do then fine. But it won't advance our understanding of what humans and machines can and can't do while making the wrong attributions (note the realist talk), can really muck up social existence for everyone.

Finally, turning to AI-aided science in particular, it is important to understand the science being done. If it is 'hypernormal science' such as molecular biology [Collins et al., 2023] then it is much more open to computerisation than a science like gravitational wave detection (though there is a way of using computers to simulate and explore the terrain of experimental potentialities there too). But that computers can aid science has been obvious since the invention of the slide rule [Collins, Thorne, 2026] distinguish the use of computers in different kinds of science.

## Thorpe and Sokolova

I owe a big debt to Charlie Thorpe. Like, I suspect, many adventurous academics I spend quite a bit of time wondering if I am the crazy one, especially when I write papers that I think are important but hardly anyone notices them. One such paper is the one on 'socialness' [Collins, 1998] which Thorpe discusses. This, paper, published in 1998, has been cited 46 times according to Google Scholar, and seven of those are self-citations. For comparison, my most cited item *Changing Order*, is slowly nearing 5000 citations, my tenth most cited paper is approaching 1000 and my twentieth is approaching 500. I'd say the socialness paper should rank in the top ten or twenty of my publications but in terms of citations it's around rank 100. Still worse, is the case of the paper called 'Interactional Imogen', published in 2020, which I consider a really crucial follow-up to the socialness one. It has been cited 11 times, of which 6 are self-citations! In it I try to work out the limits on one social group coming to understand another social group. Again, I believe it comes up with enough important insights to rank among the top 10 or 20 of my publications but is around rank 200 in terms of citations. Please read it!



So, Thorpe's discussion of 'Socialness' offers the possibility that in spite-of these works, which I consider a couple of my very best, being close to invisible in the world I work in, I may not be completely mad. And Thorpe is one the five 'proper' citers of 'Imogen'. So, you see the size of my debt!

To get to the relationship between philosophy and sociology, the socialness paper is a plea for a sociology as I (and Thorpe, and Peter Berger) understand it. This sociology goes back at least to Durkheim and his 'treat social facts as things'. As Thorpe explains, I believe humans are unique because they have 'socialness' – the ability to take part in and understand social groups whereas animals and other non-humans do not; human-like language is central to this capacity and the term 'interactional expertise' captures the ability to understand, through language, the practical face of social capacity without actually being able to execute the practices. Acquiring interactional expertise in the world of an initially unfamiliar social group is hard but, in the unread 'Imogen' paper I argue that it is possible under the right circumstances.

All this is philosophy as much or more than it is sociology, as I see it, albeit I did come up with it as a result of living with scientists during my gravitational wave study. Furthermore, if this is philosophy we are going to have to call Durkheim a philosopher. But why shouldn't we? We also call the later Wittgenstein a sociologist and agree that while Winch was calling sociology 'misbegotten philosophy' he might equally have been calling philosophy (at least, of the late Wittgenstein type which was his topic) 'misbegotten sociology'.[1] That's why I am very pleased to find Sokolova reminding me of what I wrote in 2018 after a somewhat bad tempered encounter with certain philosophers of expertise: "Is what I do here philosophy or sociology?" – "the question is misconceived".

This makes Sokolova's concern about whether philosophy is 'abandoning its birthright', as we might say, by conceding to sociology's empirical approach, very difficult to answer. We might be able to agree that philosophy has changed, but if it has, we can locate the beginning of the change with a philosopher – Wittgenstein. But, in any case, all academic subjects change and not necessarily by conceding ground to other subjects. Let us start with sociology.

[1] [Bloor, 1983] is a good place to find the claim that Wittgenstein was as much sociologist as philosopher: what he was doing was describing the way his society worked, being pretty knowledgeable about it as all citizens are without getting out of their armchairs to do empirical studies.

## Sociology has changed

What I have come to realise is that I wrote the 'socialisation' paper as a response to a change in sociology – one that I didn't fully understand at the time. But now that time has passed, and Thorpe has helped me see that this is a topic worth pursuing, the change that I was responding to was from what I still consider to be sociology's main concern, understanding the world through the concept of the social group. Sociology as practised by sociologists was changing to the defence of those who live at the interstices of social groups – those marginal to social groups; sociology's topic had become diversity. The new topic was no longer the science of the social but a political movement, often referred to as 'identity politics'. Identity politics took over from sociology as a science because the main feature of sociological science was the study of groups and the idea of groups seemed to suppress the rights of individuals who were marginal to this group or that group. The study of the old sociology was how groups worked and it concentrated on those who were full members. I'm no kind of historian but I think the driving forces behind the transformation were civil rights and gay liberation because marginality was the dominant political experience to those who were discriminated against, and the idea of social groups seemed to emphasise the marginality of the marginal. Sociology turned into the politics of those in the spotlight of discrimination.

As far as I can see, this is where sociology still finds itself and that is why we should not be surprised that sociologists are not going to read a paper which asks the question about whether one social group can learn to understand another social group and the conditions that make it is possible; they certainly aren't going to cite such a paper in their own work. The question is outlawed by the off-the-shelf political answer: marginal groups can be represented only by members of those groups. What now has to be added, though no-one seems anxious to add it, is that this means anthropology, ethnography and sociology understood as the study of unfamiliar groups by academics specially trained in understanding initially unfamiliar cultures, is impossible: at best, it is disguised colonialism.

As an aside, though the experience of marginality is real enough (my early life was lived as a marginal person), the marginal are usually as much members of social groups as anyone else, which is how they function as members of society. The social identity of being gay or being black is enormously powerful and keeps identity politics going. To understand how socialness works one needs the fractal model of society which locates socialness all the way down to the smallest groups.[2] That is the scientific concept, but it is obvious that it doesn't stop groups being

[2] For the fractal model see, for example, [Collins, 2023, Figure 1].



in conflict and smaller groups being discriminated against by dominant groups. But that's the politics, not the science.

Collins (and Thorpe) are still trying to do the older kind of sociology and the idea of 'the fractal model of society', and 'interactional expertise', and its empirical counterpart – the 'Imitation Game' – are central to it. My typical critique of artificial intelligence grows out of it. AIs are not members of social groups in the way humans are; even large language models, which are hugely successful because they are closer to be being socialised in the way humans are socialised than previous generations of AIs, fail where their socialisation departs from the human experience described by the fractal model and which our recent papers on the topic try to demonstrate.

## Becoming a philosopher?

I have never been trained in philosophy by anyone who has a philosophy degree, nor do I have any philosophical qualifications but, in continuing to do sociology in the way described above, and given what's happened to sociology, there are indications that I am now ceasing to be a sociologist in the way most sociologists understand the term nowadays, and becoming a philosopher. The idea of interactional expertise and of the Imitation Game have been criticised by sociologists because they reify the idea of the social group, but that, to me, is what sociology is about. I don't know if this was the reason, but I was astonished when our first paper, empirically exploring the idea of interactional expertise using the Imitation Game was rejected by *Social Studies of Science* (it was the first time I had ever had a paper rejected by that journal); it was then accepted by philosophy journal, *Studies in History and Philosophy of Science*![3] And, nowadays, a large proportion of my work is published in philosophy outlets, including the Imogen paper, even though, as I see it, it deals with a quintessentially sociological question – it explains the circumstances under which sociology (as well as ethnography and anthropology) can work. I have cited philosophy outlets for my work in this paper wherever I have had a choice.

It is also the case that my work is deeply rooted in philosophy. The framing concept of the book that defines my early work – *Changing Order* – is the problem of induction: I describe myself there as setting out a sociological solution to the problem of induction. And I think the philo-

[3] This paper is [Collins et al., 2006]. A comprehensive book on Imitation Games is in press (2027, forthcoming). It is called *Imitation Games and Turing Tests: New methods for exploring societies*, Cambridge, MA: MIT Press. The book has an invitation to consult an open access web source: https://github.com/igwebstore. The last item there is the letter Collins wrote to the editor of *Social Studies of Science* asking for a careful explanation of why it had been desk-rejected.

sophical roots of the idea of interactional expertise lie with Nelson Goodman's 'new riddle of induction' since Goodman claims that our grasp of practical things is found in the degree of entrenchment in the language of our descriptions of reality [Goodman, 1955]. And my philosophy of sociological reality is based on philosopher Wittgenstein. For about two decades, until he died (in 2017), I had a continual and very friendly argument with phenomenological philosopher Hubert Dreyfus, about the relationship between language and practice, Dreyfus insisting that practice was all, and me insisting that language could capture practical understanding in the absence of practice itself. Dreyfus is famous for insisting that computers could not be intelligent unless they had bodies and he wrote the definitive (1973 – 2nd edition 1992) book-length critique of AI (which, as I often explain, would have stopped my own critique of AI in its tracks if I had not been lucky enough *not to* read his book before I had worked out my own position – see [Collins, 2019]).

At this point let us turn aside to explain, if it is not already understood, that philosophy itself can have practical consequences of the kind that science produces. Dreyfus thought he was producing a result for artificial intelligence research of what I am calling. a 'scientific' type; he was telling AI researchers that they had better start working on bodies for their machines if they wanted to reproduce human-like intelligence. But Dreyfus was getting it wrong (as scientists are entitled to). What he was missing (as pointed out by computer scientist, Doug Lenat), was that the human world included people without proper bodies (the severely physically challenged) whose intelligence was unaffected. At this point the notion of 'body' in philosophers' arguments became somewhat diffuse: unfortunately, no-one took responsibility for describing the limits to what counts as a body such as would render the need for body in an intelligent being open to falsification.

I think the 'Imogen' paper brings the practical implications of the Dreyfus-like philosophical argument back via the table of levels of AI, reproduced in, for example, [Collins, 2021, Table 1]. The lower levels of the table are concerned with individual AIs but Level 5 is autonomous societies of AIs. Imagine some thousands of *tabula rasa* AIs – AIs without any cultural experience – dropped onto a desert island with conditions for survival provided (the kind of experiment we might once have envisaged with newborn infants instead of AIs in order to resolve the nature/nurture problem). It seems obvious that for such an artificial society to develop human-like language the machines would need human-like bodies. To make the point as obvious as possible, if they didn't have bodies capable of playing tennis, they would never develop tennis-related language nor the capacity to understand it when spoken by a human visitor to their society. But this seems wrong for a single *tabula rasa* AI dropped into an already functioning human society – indeed, existing LLMs can already handle tennis-related language (I am not saying they



*understand* what they are talking about but, as Filatova points out, the talk is fluent enough for us not to worry about their understanding when we engage with them unless we are trying to investigate their abilities). So, the practical implication of this paragraph's-worth of philosophy is 'don't worry about the body when dealing with AIs dropped into human societies but do worry about bodies when thinking about autonomous societies of AIs'. A lot of this kind of thing is worked out in the 'Imogen' paper (published in a philosophy journal, by the way).

## Formative aspirations

But there is another problem: not every 'philosopher' is really a philosopher! To worry about Sokolova's problem – is philosophy losing its special warrant in taking on sociology's empirical practices? – we have to know what philosophy really is. I'm pretty sure that it is not the sum total of the activities of all those who are referred to as 'philosophers' (though it has become very hard to say such things since constructivist sociology of science began to flourish). For example, I know there are a lot of highly trained and highly respected philosophers whose work involves their becoming experts on the writings of other philosophers and publishing the results. I think this is philosophical scholarship not philosophy. On the other hand, I am pretty sure that the later Wittgenstein is a philosopher even though David Bloor describes him as a sociologist – Bloor's not wrong it is just that in his work philosophy and sociology are not distinguishable.

So, what is philosophy? What are the 'formative aspirations' of a philosopher.[4] I think that is a hard question and I really hope philosophers will have a crack at it. I think the nearest I can get is the old story about the sage and novice who asks what supports the Earth (the Earth being a proxy for truth). The question-and-answer session in which the novice continually asks 'but what does that stand on', goes through the shoulders of Atlas, via an elephant, to a turtle, culminating in the sage saying 'from there on its turtles all the way down'. I think philosophy's formative intention is to get to the bottom turtle in so far as it is possible. And that, it seems to me, is where sociology and philosophy bump into each other because the sociology of knowledge is a very philosophical idea that is as

[4] This uncomfortable question relates to what I have been doing in trying to find a definition for 'science'. Whatever that definition is, it does not include the activities of cheats and liars who work in the scientific profession, it refers only to those who conform to the 'formative aspirations' of science as set out in, for example, [Collins, 2023, Table 1], or for a slightly improved version which will form the basis for future work see Collins and [Evans, 2026, Figure 1.2]. For the origin of the notion of 'formative aspiration', or 'formative intention', see [Collins, Kusch, 1998].

terrifying as philosophical scepticism – we only know what we think we know because of how we have been socialized.

I think that before Sokolova worries too much about whether philosophy is losing its special warrant she needs to find a way to define philosophy and that is itself a terrifyingly difficult question.

It seems to me that Thorpe and I are in a position to answer the question in respect of sociology if we are brave enough: sociology's formative aspiration is to understand the special nature of human social groups. I say this requires us to be brave because it excludes what most sociologists actually do these days. Maybe we should keep our mouths shut! But for me, I think the collision of sociology and philosophy at the level of the bottommost turtle we can find explains why I still don't know whether I'm a sociologist or a philosopher but I would say that Sokolova should not worry too much because in terms of what sociologists and philosophers do, philosophy is in the better shape.

## Afterword

I consider, above, that I have become a philosopher and I separate myself from the majority of sociologists, whose academic life is primarily driven by the politics of diversity. But in recent years, as result of political events in 'the West', my work has also become primarily driven by politics. What I am trying to do these days is reform science and science-related education in a way that will inform the citizenry of the need for truth in politics if democracy is to be preserved and in trying to explain the role of science as a bulwark against the erosion of truth. The political battle to be fought is not for the preservation of underprivileged minorities (though this remains a worthy aim) but for the preservation of expert elites against the predations of would-be dictators. One does not always know where an academic career will take one.[5]

## References

Bloor, 1983 – Bloor, D. *Wittgenstein: A Social Theory of Knowledge*. London: Macmillan, 1983.

Collins, 1998 – Collins, H. "Socialness and the Undersocialised Conception of Society", *Science, Technology and Human Values*, 1998, vol. 23, no. 4, pp. 494–516.

Collins, 2018 – Collins, H. "Studies of Expertise and Experience", *Topoi*, 2018, vol. 37, pp. 67–77. DOI: 10.1007/s11245-016-9412-1.

[5] The latest expression of this is [Collins, Evans, 2026], which has the title, *Veritocracy: Truth, science, and how to preserve democracy,* but [Collins, Evans, 2017], *Why Democracies need Science,* is a predecessor.



Collins, 2019 – Collins, H. "Remembering Bert Dreyfus", *AI and Society*, 2019, vol. 32, no. 2, pp. 373–376.

Collins, 2020 – Collins, H. "Interactional Imogen: Language, Practice and the Body", *Phenomenology and the Cognitive Sciences*, 2020, vol. 19, no. 5, pp. 933–960. DOI: 10.1007/s11097-020-09679-x.

Collins, 2021 – Collins, H. "The Science of Artificial Intelligence and Its Critics", *Interdisciplinary Science Reviews*, 2021, vol. 46, no. 1–2, pp. 53–70. DOI: 10.1080/03080188.2020.1840821.

Collins, 2023 – Collins, H. "Science as a Counter to the Erosion of Truth in Society," *Synthese*, 2023, vol. 202, no. 145, pp. 1–23. DOI: https://doi.org/10.1007/s11229-023-04357-2.

Collins, Evans, 2017 – Collins, H., Evans, R. *Why Democracies Need Science*. Cambridge: Polity Press, 2017.

Collins, Evans, 2026 – Collins, H., Evans, R. *Veritocracy: Truth, Science, and How to Preserve Democracy*. New York, NY: Bloomsbury Academic, 2026.

Collins et al., 2026 – Collins, H., Hall, M., Evans, R., Weinel, M. *Imitation Games and Turing Tests: New Methods for Exploring Societies*. Cambridge, MA: MIT Press, 2026 (forthcoming).

Collins, Kusch, 1998 – Collins, H., Kusch, M. *The Shape of Actions: What Humans and Machines Can Do*. Cambridge, MA: MIT Press, 1998.

Collins et al., 2006 – Collins, H.M., Evans, R., Ribeiro, R., Hall, M. "Experiments with Interactional Expertise", *Studies in History and Philosophy of Science*, 2006, vol. 37A, no. 4, pp. 656–674.

Collins et al., 2023 – Collins, H., Shrager, J., Bartlett, A., Conley, S., Hale, R., Evans, R. "Hypernormal Science and Its Significance", *Perspectives on Science*, 2023, vol. 31, no. 2, pp. 262–292.

Collins, Thorne, 2026 – Collins, H., Thorne, S. "Large Language Models and Scientific Discourse: Where's the Intelligence?", *Synthese*, 2026 (accepted, February 2026). DOI: https://doi.org/10.1007/s11229-026-05531-y. Open access: https://rdcu.be/fbpKm. Preprint: *arXiv:2603.23543*.

Collins et al., 2026a – Collins, H., Grote, H., Newbury, P., Sutton, P., Thorne, S. "Large Language Models Eroding Science Understanding: An Experimental Study", *arXiv preprint arXiv:2604.25639*, 2026 (under revision for resubmission to *AI and Ethics*).

Collins et al., 2026b – Collins, H., Thorne, S., Newbury, P., Evans, R. "More Problems of Large Language Models in Comparison with Human Knowledge", *AI & Society*, 2026. DOI: http://doi.org/10.1007/s00146-026-03021-w.

Collins, Yearley, 1992 – Collins, H.M., Yearley, S. "Epistemological Chicken", in Pickering, A. (ed.), *Science as Practice and Culture*. Chicago: University of Chicago Press, 1992, pp. 301–326.

Dreyfus, 1972 – Dreyfus, H.L. *What Computers Can't Do*. New York: Harper and Row, 1972 (2nd edition: *What Computers Still Can't Do*, 1992).

Goodman, 1955 – Goodman, N. *Fact, Fiction, and Forecast*. Harvard University Press, 1955 (1st edition).

Weizenbaum, 1976 – Weizenbaum, J. *Computer Power and Human Reason: From Judgement to Calculation*. San Francisco: W.H. Freeman, 1976.