Humans Not Instruments

Author(s): Harry Collins

Humans Not Instruments*

Harry Collins†

I argue that it is serious mistake to treat instruments as having parity with humans in the making of scientific knowledge. I try to show why the parity view is misplaced by beginning with the "Extended Mind" thesis which can be seen as an individualistic version of Actor/ant Network Theory, and then move on to instruments. The idea of parity cannot be maintained in the face of close examination of actions as simple as doing a calculation or accepting the reading of an instrument. The key difference is that humans are embedded in language communities–the locus of knowledge-making–and nothing else is.

I understand there is a movement in the social studies of science which treats instruments as intelligent actors or participants in a social knowledge. It would fit well with tendencies such as the very popular "Actor/ant Network Theory" (ANT).[1] ANT treats non-humans, such as instruments, as equivalent to humans. But to treat humans and non-humans symmetrically is to labour under a deep misunderstanding of the way the world works. I will try to explain.[2]

It is impossible to understand what goes on in a laboratory unless a clear distinction is maintained between humans and non-humans.

---

* Received 12 January 2010.

† Harry Collins is Distinguished Research Professor and Director of the Centre for the Study of Knowledge Expertise and Science (KES) at Cardiff University. He has books in press on tacit knowledge and on discovery in physics and a book in draft on a new method for social science research–the Imitation Game. Much of his recent work can be followed via www.cf.ac.uk/socsi/expertise.

[1] This approach is most closely associated with Bruno Latour (2007) and Michel Callon (1986). I refer to it as actant' network theory since the idea of an 'actor' in social science involves intention whereas the entities belonging to ANT, which draws on the metaphor of literary criticism, are not humans but simply nodes in a network of relations. The qualities of these nodes are just what is ascribed to them by the network and it would be quite inappropriate to think of them as having intentions or understandings of the kind that are taken to be the basis of interpretative sociology.

[2] I have resisted this approach from the outset (Collins and Yearley 1992; Collins and Kusch 1998 and passing references in many other publications). Certain groups of philosophers of science have, perhaps, been encouraged to treat this tendency as a still more dominant feature of social studies of science than it is as a result of the central place given to Latour's work in Ian Hacking's widely read *Social Construction of What*? (2002).

Humans have language which supports social life and non-humans do not. Humans argue about the output of non-humans and publish the results while non-humans do not. Humans are the interpreters of what non-humans do while non-humans are not the interpreters of what humans do. Non-humans cannot be interpreters because interpretation is a collective activity and non-humans do not have the means–language–to engage in collective interpretation. See Collins (2010) and many items at www.cf.ac.uk/socsi/expertise for a discussion of the role of language in human understanding. This seems obvious. How has the mistake of treating as equivalent humans and non-humans happened?

THE EXTENDED MIND

Rather than turning back to ANT, which has already been criticised at length (see footnote 2), let us look at a particularly persuasive version of the idea that humans and other entities are part of the same collective: Andy Clark's idea of the "extended mind" is in many ways homologous with ANT in the way it takes humans and non-humans to work together[3] (Clarke and Chalmers 1998; Clark 2002). If I can sum up the thesis briefly without caricature, it holds that our cognitive abilities are extended by the material world around us and that, therefore, when we do something like a calculation with the assistance of paper and pencil, or a calculator, the entity doing the calculation should be not be thought of as the brain or mind but the entire ensemble of whatever is inside us and whatever is outside us that makes a contribution to the result. Clark, who has a nice way with words, tells us that our concept of the human should not be bounded by skin and skull. Particularly engaging is the willingness of Clark, in cooperation with David Chalmers, to push this idea into the body as well as outside it. Thus, the pencil and paper, or calculator, are treated as being on a par with those physical features of the brain that are engaged in a calculation. In sum, if there is a physical counterpart to the "mind" that does a calculation, it is all the bits of stuff both inside and outside the body which fired up during the course of the action; there is no reason at all to treat the boundary between inside and outside as anything special. In fact, Clark wants to treat the mind as an ensemble of physical entities both inside and outside the "skinbag" (a less felicitous choice of term).

The extended mind idea is attractive and persuasive. As can, I think, be seen, it is an individual-based version of the ANT view of the world in so far as they are both informed by the idea that there is no significant

[3] For an earlier debate concerning the asymmetry discussed here see Collins, Clark and Shrager (2008).

distinction between that which is human and that which is not.[4]

My problem with the extended mind is similar to my problem with ANT. While it is the case that the non-human elements in a mind-network or a social network contribute to the functioning of the network they do not contribute in a way which is symmetrical with the human elements. Thus, if I say to an ANT-enthusiast, as I often do, "my computer/dog/plough is not social" they respond along the following lines: "but look at the extent to which the computer/dog/plough is integral to your social life. How can you possibly understand social life without taking into account how different it would be without your computer/dog/plough?" The answer is that I cannot understand social life without taking into account my computer/dog/plough, but that does not make them social beings. Thus, one might say, "look how different your social life would be without your heart/oxygen/the sun, but that does not make my heart/oxygen/the sun into social beings." To be consistent, however, that is where you have to take the logic.[5]

Let us try to reveal the difference between the human and non-human contribution by jumping off from the extremely plausible extended mind. Let us suppose I want to do the following calculation: I want to convert my height in inches to my height in centimetres. My height is 69 inches and there are 2.54 centimetres to the inch.[6] "I" can choose to do it in at least three ways: I can use a paper and pencil; I can use a slide-rule; and I can use a calculator. As far as I can see, the paper and pencil, the slide rule, and the calculator cannot make the corresponding choices. They do not, for example, have a choice about whether to include me in the process of calculation. This isn't a point about free will; it is a point about what can be substituted for what in the process of converting a height in inches to a height in centimetres. If I don't participate but just leave the paper and pencil, slide-rule, and calculator to get on with it there will be no calculation while any two of the other three can be left out.

The asymmetry does not imply that the mind does not extend itself

---

[4] Latour believes he has developed a new ontology while my understanding is that Clark is happy to treat his view as a heuristic–something that is an aid to uncovering new things that might otherwise stay hidden (Andy Clark, January 2010, communication with author). If I have understood the matter correctly, Clark's attitude resonates with "methodological relativism" in sociology of scientific knowledge (e.g., Collins [1985] 1992), where relativism is taken not to be a philosophical position but a methodological approach.

[5] To be fair to Latour, he is consistent on this point; for him, your heart, oxygen, and the sun are social/non-social beings. Latour quite self-consciously "deconstructs" the distinction between the social and the non-social thus destroying the use of the social as something with causal efficacy. It's a self-consistent, even elegant, position but, unfortunately, it's just wrong.

[6] This example is taken from Collins (1990) and developed a little further in this paper.

into these non-material objects but it does show that there is no parity between the elements that go to make up the mind–the ensemble which comprises the mind seems to have some non-substitutable stuff and some substitutable stuff. It is very hard to get a grip on what the non-substitutable stuff is but I am going to try to bring things back to ANT by taking one further step, which, I believe, shows that the non-substitutable stuff is the thing that has social life via its language.

CALCULATING MY HEIGHT

Suppose I choose to use the calculator to convert my height from inches to centimetres. The answer I will get is 175.26 centimetres. But a human being who is embedded into British social life via language will know that this answer is wrong. If I am that person, though I know the answer is wrong, I do not know what would be right. I only know the right answer might be 175 cm, it might be 175.3 cm, or it might be 175.5 cm. It won't be 175.26 because I know that it is a convention in my society that heights are never given to that level of accuracy–people say it does not make sense to measure height to one hundredth of a millimetre because height varies much more than that between breaths; part of the art of calculation is knowing what counts as an appropriate level of approximation.[7] Just what level of accuracy is conventionally accepted when heights are given in centimetres in Britain I do not know because I have spent almost my entire life discussing heights in inches and am not embedded in any of those groups (such as the medical profession) which measures heights in centimetres; that is why I do not know which of the three approximations is counted as correct. The only way I can find out is to embed myself in such groups and discover their conventions (or in this case, just ask someone who is so embedded).

Now one could say that this problem is easy to solve–one just asks the person who is embedded for the correct convention and adds some extra keys to the calculator such as one that says "human height calculation" with the appropriate level of approximation linked to the key via the program. This is the solution known in artificial intelligence circles as the provision of "scripts." One can see why it will not work: (a) there are going to be an awful lot of extra keys (or extra sub-routines in the case of serious artificial intelligence; (b) there remains the problem that the human and only the human can say what the right script is in the lived context; and (c) the scripts are in a state of continual flux depending on how human social life plays out and no-one can predict how human social life will play out. That is why artificial intelligence does not work and will not work until we

---

[7] See Kuhn (1961) for a brilliant essay relating to this point.

find a way of embedding a computer into social life–and we have absolutely no idea how to do it except giving an artificial intelligence a human minder.

Note the asymmetry: there is nothing that can be substituted for that human minder because only the human minder is connected into society and we do not have anything else that can do the job. The reason there remains an "I", even for the extended mind hypothesis, is that only an "I" is connected, via language, into society and it is society that decides what the right answer to things is. (It should be a trivial matter for the reader to explain why modern spell-checkers make suggestions about word corrections rather than automatically doing the corrections–except in trivial cases. Even as simple an "artificial intelligence" as a spell-checker needs a human minder.)[8]

ANT is much more ambitious than the extended mind. ANT has the ambition to explain why "this" feature of social life is present and "that" is absent. It tries to explain it by assembling all the forces, human and non-human, that contribute to some state of affairs. But all the forces you can think of conspire toward there being spell-checkers, speech transcribers, and the like that do not need human minders–build one and you will be richer than Croesus! ANT simply cannot explain why they are not here with us now and ANT will never explain it because in insisting on the symmetry of the human and non-human it misses the entire point about the centrality of language–it misses the entire point of what it is to be social and how unique to humans being social is. When it comes to these things, ANT, plainly and simply, gives the wrong answers.[9]

[8] What do I mean by a trivial case of spelling? Something like the following misspelling: "becuase." Damn! It's not so trivial after all. The spell-checker corrected it automatically first time (try it) and I had to type the middle letters again to (mis?)spell it for this note. Even here a human minder is needed: there are no trivial cases except where live context allows them to be trivial and live context can't be recognised without a human minder.

[9] Some relatively promising approaches to AI research attempt to keep in touch with social life via the flux of linguistic interaction on the internet. I doubt that these approaches will work: we need to ask ourselves, could an anthropologist come to understand the social life of a strange people solely by interacting with them via the internet? In physics we know that what counts as credible can only be understood from immersion in the oral culture, not the written culture (Collins 2004; see also the work by Weinel on the knowledge of those who accepted Thabo Mbeki's views on HIV transmission which can be found at www.cf.ac.uk/socsi/expertise). Nevertheless, whether the "internet approach" would work or not, the very attempt to make it work acknowledges that the problem of linking into social life has to be solved if artificial intelligences are to substitute for us humans. For ANT everything is already linked into social life so there is no problem to be solved. That is why it goes so wrong.

INSTRUMENTS AND INSCRIPTIONS

Let us apply the above arguments to scientific instruments. For ANT, an instrument is taken to be an "inscription device" that gives rise to an "immutable mobile"–data on paper or other media that can travel around without changing–and this is what gives it its power.[10] An instrument/inscription device would have no force without the idea of immutable mobile because if inscriptions were entirely mutable they would be mere scribbles and an instrument that scribbles can do no persuasive work; the immutability of inscriptions–that is, their resistance to indefinite interpretation–is what determines the usefulness of this set of ideas. The proper subject in the case of inscription devices and immutable mobiles is, then, the degree of mutability (or interpretative flexibility).[11] How mutable are inscriptions?

Here again, the degree of mutability and the range of interpretation is, to a very large extent, in the hands of the social group upon which the inscriptions are inflicted. At the most basic level, this is likely to vary from science to science and from phase of science to phase of science. I have the strong impression that in the biological sciences, observation of which formed the impetus for the first formulation of the inscription device and immutable mobile ideas, and in the particular phase of biological science that was being observed, there was, as a matter of fact, little mutability and little interpretation and this is what started things off on the wrong foot. But lack of mutability is certainly not always present in the case of the biological sciences. My colleagues working in genomics have shown me quotations that make it clear that results coming from other laboratories are often ignored unless there is a personal relationship or some reassurance based on anecdotal evidence that leads them to be taken seriously. It is not what happens *within* a laboratory that matters, it is what happens *between* laboratories. And, as I know from my own research, mutability is very much not to be found at the frontiers of physical science.

Figure 1 is an example of the "inscription" produced by an early gravitational wave detector. Now you could say, with a kind of deliberate naïveté, that this inscription is the same for every scientist who sees it and, like an intelligent actor, it simply gives you an interpretation that has to be taken into account, but you would not be saying very much. The crucial question is whether the inscription is just noise or something of value and that depends not on the paper and ink but on the readers' understanding

---

[10] These ideas were present in *Laboratory Life* (Latour and Woolgar 1979) even before full-blown ANT came along.

[11] The term "interpretative flexibility" is taken from Collins (1981); see also Collins ([1985] 1992).
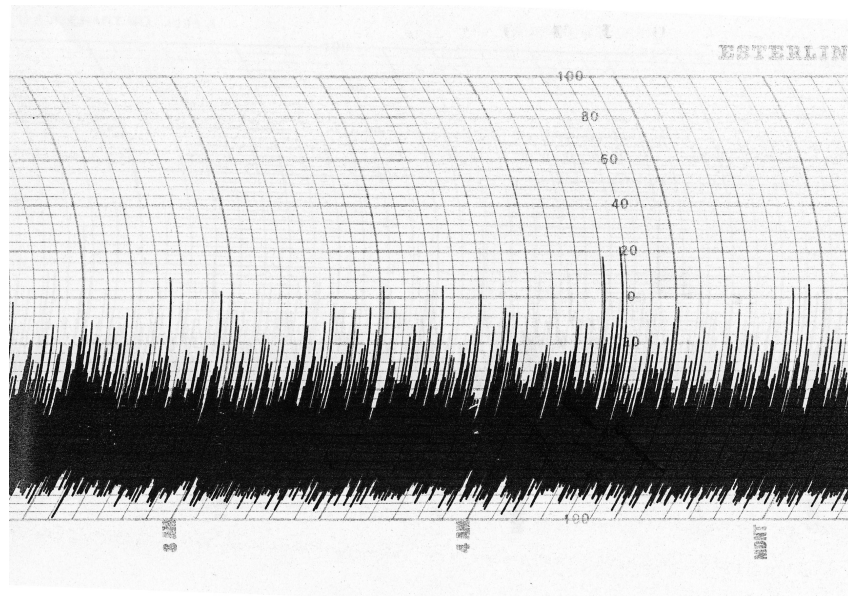
Figure 1. An example of a trace from an early gravitational wave detector (Collins 2004, 79).

of how it was produced and how much they trust its producer; this understanding comes before they start to work out ways of extracting any lurking signal from the noise.[12] As it happens, the state of play in gravitational wave physics is such that it is now considered impossible for any inscription produced before 2003–though many were produced from the late 1960s onward–to contain a meaningful gravitational wave signal. Indeed, Collins (2004) discusses a paper published in 1996 which contains the inscription (or interpretation of inscriptions) represented in Figure 2.

It turns out, however, that this inscription was further "interpreted" by default in that no-one read it (outside of the author, the journal editor, and Collins). The 1996 paper reported results of certain early gravitational wave detectors and, by 1996, the human-led consensus was that these were to be counted as worth no more than scribbles–not even worthy of criticism or even a cursory reading.

Talking about instruments as devices that produce inscriptions is, then, a misleading way of putting things because it suggests that it is easy to arrange for something stable and constant to travel between communities of scientists. The real problem is how any inscription can be stable in any

---

[12]See Collins (2011) for an illustration of how a signal can be extracted from an inscription of this kind. Collins (2011) is a *book length* treatment of a two-year process in the course of which a community interpreted one particular inscription–very similar to that shown in Figure 1–produced at one particular time on one particular day.

Summary. — Gravitational-radiation-antenna outputs were plotted, for 20 s periods which included one gamma-ray burst trigger time. Data were obtained for two gravitational-radiation antennas. For the first 80 burst times which were studied, it is observed that either the largest or second-largest gravitational-radiation detector pulse peak occurs within 0.5 s of the gamma-ray burst trigger time for 20 of the 80 gamma-ray bursts, for the larger of the two gravitational antennas. The probability that the correlations are accidental is estimated as approximately $6 \cdot 10^{-5}$.

PACS 95.30Sf – Relativity and gravitation.

Figure 2. Inscription from a paper published in 1996 (Weber and Radak 1996, 687; for discussion see Collins 2004, 366-68).

way under the heat of varying interpretations. Nearly all the inscriptions that I have studied have been dominated by their mutability. As soon as they are out of the hands of their producers everyone starts doubting their meaning or trying to establish it. But even in non-disputed subjects inscriptions are not stable and even the relatively unqualified can question their meaning. I remember pointing to a graph in a book about cosmic background radiation–a subject about which I know little–and saying to one of the scientists who was involved in the original research project, "That line fits the data points so exactly that I don't believe it." The scientist told me I was right to be suspicious. What was immutable about that chart?

To understand still more about what an inscription is, one might reflect again on Figure 1. How did you know when you first looked at it that it was what I said it was? What is interesting about that inscription is not the ink on the paper but the body of institutions and your socialization into them that led you to take it to be what I said it was without even a pause. That has to happen even before we even get to the overwhelming mutability of what the ink marks mean. The ink marks? They might be interesting to a scientist well-embedded into some network of trust (or in that particular case, more like lack of trust) but there isn't much interesting about them to a social scientist. There isn't much of interest to a social scientist because all the interpretation we have discussed is done by humans and every time we have mentioned consensus it is formed by humans; the non-humans are not doing any interpreting and they cannot contribute to the consensus because they don't have the language to do it.

THE LIFE OF INSTRUMENTS

Thus, the entire previous section has been an exhibition of the asymmetry of humans and non-humans. Humans are everywhere–there are disagreeing gravitational wave scientists, there is me refusing to accept

a graph showing the output of a cosmic background detection experiment, and there is you wondering about why you believe Figure 1 is what I say it is. Now, of course, the inscriptions we looked at could be different, and this might make one interpretation harder to hold than another, but even though it is harder to cycle to work than to drive, this does not make my car or my bike interpreters of my life, anymore than the sun is an interpreter of my life.[13]

It is not logically incoherent to think of the world of non-humans as invested with human-like qualities; religious leaders and storytellers have done this through the ages and some of us spend a lot of time complaining about those who treat their cars as extensions of their personality. It is also not inconceivable that thinking of things this way might occasionally be of heuristic value in the history and philosophy of science. I have to say that I have never discovered this value. On the contrary, I have always found it misleading and, where applied to fields I understand to some extent, such as artificial intelligence and physics, positively harmful.[14] In this it contrasts markedly with the usefulness of, say, "methodological relativism": one gives the wrong answers, one gives the right answers. To the extent that one wants to avoid mistakes, it is important to remember the difference between the human and the non-human and to remember that it is only humans who interpret what the outputs of instruments mean. Never treat an instrument as an intelligent actor or a participant in social knowledge unless you have a very good methodological reason for doing it–a reason that one could and should explain clearly.

HARRY COLLINS
School of Social Sciences
Cardiff University
Cardiff CF10 3WT
UK
CollinsHM@cf.ac.uk

[13] For further discussion of this difficult point using the term "affordance" see Collins (2010).

[14] It is also positively harmful in another area of which I have a passing acquaintance: sports umpiring. In international tennis, line calls have been "delegated" (a word drawn from the repertoire of ANT) to technologies such as "Hawk-Eye." It is extremely dangerous to do this, however, without considerable thought about the different effects that delegation has in respect of different devices and different kinds of judgement. For example, in the case of close line-calls in tennis Hawk-Eye has been given what we might call "ontological authority"–the right to make truth–whereas when the call is very close the machine's "decision" is close to random. The entities that have to decide if this is appropriate are humans. To set a precedent by treating Hawk-Eye as having parity with human judgement could be a step toward a technological dystopia.

REFERENCES

Clark, Andy. 2002. *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence.* Oxford: Oxford University Press.

Clark, Andy, and David Chalmers. 1998. The Extended Mind. *Analysis* 58: 7-19.

Collins, Harry M. 1981. Stages in the Empirical Programme of Relativism. *Social Studies of Science* 11: 3-10.

Collins, Harry M. [1985] 1992. *Changing Order: Replication and Induction in Scientific Practice.* Beverley Hills: Sage Publications. Rev. ed. Chicago: University of Chicago Press.

Collins, Harry M. 1990. *Artificial Experts: Social Knowledge and Intelligent Machines.* Cambridge, MA: MIT press.

Collins, Harry M. 2004. *Gravity's Shadow: The Search for Gravitational Waves.* Chicago: University of Chicago Press.

Collins, Harry M. 2010. *Tacit and Explicit Knowledge.* Chicago: University of Chicago Press.

Collins, Harry M. 2011. *Gravity's Ghost: The Equinox Event and Science in the 21st Century.* Chicago: University of Chicago Press.

Collins, Harry, Andy Clark, and Jeff Shrager. 2008. Keeping the Collectivity in Mind? *Phenomenology and the Cognitive Sciences* 7 (3): 353-74.

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

Collins, Harry, and Steven Yearley. 1992. Epistemological Chicken. In *Science as Practice and Culture,* ed. Andrew Pickering, 301-26. Chicago: University of Chicago Press.

Callon, Michel. 1986. Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay. In *Power, Action and Belief: a New Sociology of Knowledge?* ed. J. Law, 196-233. London: Routledge and Kegan Paul.

Hacking, Ian. 2000. *The social construction of what?* Cambridge, MA: Harvard University Press.

Kuhn, Thomas. 1961. The Function of Measurement in Modern Physical Science. *ISIS* 52: 162-76.

Latour, Bruno. 2007. *Reassembling the Social: An Introduction to Actor-Network-Theory.* Oxford: Oxford University Press.

Latour, Bruno, and Steve Woolgar. 1979. *Laboratory Life: The Social Construction of Scientific Facts.* Beverly Hills: Sage Publications.

Weber, Joseph, and B. Radak. 1996. Search for Correlations of Gamma-Ray Bursts with Gravitational-Radiation Antenna pulses. *Nuovo Cimento B* 111 (6): 687–92.