SYMBOLS, STRINGS AND SOCIAL CARTESIANISM: RESPONSE TO MIHÁLY HÉDER

HARRY COLLINS

Centre for the Study of Knowledge Expertise and Science
School of Social Sciences, Cardiff University, Cardiff CF10 3WT, UK
CollinsHM@cf.ac.uk

ABSTRACT

I respond to Héder's analytical critique of *Tacit and Explicit Knowledge*. Héder concentrates on the first part of the book – the way the idea of strings is used to clarify the notion of explicit knowledge. I find his critique encourages me to think there may be more to the idea of strings than just a foundation for the second part of the book; the notion of explicit knowledge really does need much more careful analysis and it should not be confounded by the mixed-up notion of 'symbols'. Unfortunately, Héder does not always distinguish between different levels of philosophical analysis: the skilled philosopher can doubt anything and everything but if we are to use philosophy to clarify things of substance we must narrow the critique and take most of the world for granted – we must be parsimonious about what we choose what to make strange. Illuminatingly, Héder show that some of my distinctions between conditions of communication are not what I thought they were. The distinction between enabling communication with longer strings and with the addition of hardware is not as clear as I thought it was. Indeed, when it comes to computers, the same string enhancement can be accomplished sometimes with hardware and sometimes with software. My solution to this problem is to refer back to general usage within the form-of-life of the transmitters of strings. I thought I was ignoring transmitters in my analysis but it turns out I need them. I thoroughly disagree with Héder's invocation of a new ontological entity, the 'pattern processing system'. Computers should be thought of as continuous with cause and effect in the natural world whereas human knowledge depends on language which computers do not possess. There are only two kinds of entities, those that can accomplish polimorphic actions and those that cannot.

Keywords: explicit knowledge; tacit knowledge; strings; symbols; ontology; polimorphic actions.

1. INTRODUCTION: THE TWO HALVES OF TEK

'[I]t is undeniable that Collins brought a brand new approach into the debate over tacit and explicit.' (Héder 2012:53) Truly one is lucky to have such generous critics and I thank Mihály Héder for considering my book Tacit and Explicit Knowledge (TEK) so carefully and provocatively. There are a number of points in his comments that have led me to think about things in new ways. There are some others that I

don't quite understand, the problem exacerbated by his tendency to ask questions and then answer them himself. There are also some criticisms that seem wrong. The balance is very much on the side of the positive and constructive.

My first paper discussing tacit knowledge was published in 1974 and I have been writing about it, on and off, ever since. When I set out to write TEK my idea was that I would draw everything I knew about tacit knowledge together in one place. I thought it would take me about four weeks to write the first draft. But what happened is that I found I could not write the book without starting with explicit knowledge and I found I did not understand explicit knowledge. That is why there are two halves to the book, the first dealing with explicit knowledge and the second with tacit knowledge. The first draft of the book took me two years to write rather than four weeks and the majority of that time I was working on the first half. But I still feel that the first half is raw, insecure and incomplete. The more I think about explicit knowledge the more remarkable and mysterious does it seem and this puzzlement has continued to grow since the book was published: Here is a person A, who cannot do some task X. A person B then causes A to be impacted by something physical – let us say some air vibrations. The air vibrations have no obvious causal connection with X – X is one thing, air vibrations are something completely different. But after the impact A can do X. That's weird!

The rawness and incompleteness did not worry me too much when I finally finished writing TEK. As far as I was concerned, the point of the first part of TEK was to find ways of talking about the explicit that would make it possible to set out the second half – the three-way classification of tacit knowledge. I saw the first part as a ground clearing exercise that would make the space for the second part. I knew that the first part might be doing little more than displaying my ignorance and reviewers might tell me that the meaning of explicit had been better dealt with by earlier thinkers but even if that happened I was content that the first part had done its job of setting up the second part. Time is drifting by, however, and I still do not know where else to look to find the resolution of the problem of explicit knowledge. I am beginning to wonder if the 'string' business is worth something in itself. So it is nice that Héder (2012) chooses to concentrate on explicit rather than the tacit knowledge aspect of TEK and especially nice that he says the approach is original.

2. POLANYI AND STRINGS

I will begin with Héder's contrast between TEK and Polanyi's *Personal Knowledge* (PK) which seems to me to be apt and which takes us straight to strings. Thus, I will develop a little further something that appears in TEK as little more than a footnote. Following Héder, it seems that TEK and PK deal with two different problems. TEK deals with the nature of knowledge, PK deals with the nature of *human*

knowledge. The difference arises, perhaps (something suggested in TEK), because of the explosion in artificial intelligence that characterised the second half of the Twentieth Century. As a sociologist of knowledge I found the 'hype' surrounding artificial intelligence could not be escaped and I found myself writing two books critiquing it (Collins 1990; 1998). Subsequently, for me, central to the problem of understanding knowledge, is the contrast between what machines can do and what humans can do.[1] I am guessing that this was not a central concern for Polanyi and that is why he concentrated on humans rather than on knowledge.

If you are primarily interested in humans it is reasonable to ask how much of their knowledge can be captured in *symbols*. It is an empirical question – you can look around and see how much is being done with symbols and how much is not being done with symbols. But, if you are interested in knowledge, that is a bad place from which to start. The reason is that the term 'symbol', and its bedfellows, 'sign', 'icon', etc., have a strong element of circularity. A 'symbol' is something that carries meaning to humans. We already know that if we are dealing with symbols they are meaningful. That's why Polanyians are always having to point out that the meaning of symbols is not provided by other symbols. Meaningfulness is unaccountable: it is tacit and that is why explicit knowledge – symbols – rest on tacit knowledge. But the same symbol is not a symbol when offered up to a machine – machines do not deal in meaning. So machines have to be dealt with quite differently to humans if you start with symbols.

It seems to me that starting with strings sweeps away many of these confusions. Sometimes a string is a symbol and sometimes it is not. The question then becomes clear – when is a symbol a symbol and when it is it not a symbol and only a string? And what can strings do when you present them to humans and machines. It becomes immediately obvious, for example, that sometimes the same string can act on a human in the way a string acts on a machine – which is what nearly everyone misses – and sometimes it can act as a symbol – something with meaning. The difference is worked out on page 17-18 of TEK with the example of the sergeant-major. This kind of thing is just a mess if you start with symbols.

Furthermore, the mysteriousness of the efficacy of strings on humans when they act in a meaningful way is much more striking than the efficacy of symbols because symbols are already efficacious by definition. More and more, it seems to me to be correct that if you want to deal with knowledge and its transmission without getting mixed up, it is best to start with strings. That may be the most important contribution of the first half of TEK.

---

[1] This is very much in the tradition of Hubert Dreyfus (e.g. 1972) though nowadays he and I have diverged markedly (e.g. Selinger et. all 2007). This is especially notable in respect of the relationship of language and practice but also in the sociological rather than individualistic approach.

The thing that strings do for you is render the familiar strange. I am not a trained philosopher so this is hubris, but I think that making the familiar strange is the fundamental move in all of philosophy. Symbols do not seem strange – we use them all the time – but when we notice that transmitting a symbol is 'really' transferring a string then we notice what a strange thing is going on.[2] I would like to think this kind of thing is analogous to Hume noticing how strange it is to reason from induction; I would like to think that the superficial but 'soothing' resolution provided by the notion of cause is analogous to the superficial but soothing resolution provided by the notion of symbol. In both cases the word helps us not notice that something very odd is going on.

As a sociologist I have tried to show that the problem of induction presents itself in the day-to-day practice of scientific discovery: how do scientists get to conclude that certain experimental results are regularly repeatable and others are not? Likewise with strings: for the sociologist, rendering explicit knowledge strange is not the end point but the starting point for questions about how knowledge and abilities are transferred in day-to-day life.

I think you have to choose what things to be puzzled about and what not to be puzzled about given that the skilled philosopher can, or ought to be able to, make anything seem puzzling. Héder has demonstrated a lot of skill in making things puzzling. The trouble is that some of the puzzles he poses are of a very general nature – general problems of philosophy – rather than puzzles that relate to the particular problem of tacit and explicit knowledge. For example, consider his question about whether adding computer chips to a plastic plate is giving new abilities to the plate. This is a good philosophical question but one which refers to the general problem of the identity of things, not the problems dealt with by TEK. For example, imagine I am at the Hungaroring talking to a McLaren engineer. He says 'I tune the engine to make the car go as fast as possible'. Lewis Hamilton walks past and says 'no, you tune the engine to make *me* go as fast as possible.' I say 'no, you tune the engine to make that speck of dirt on the exhaust pipe go as fast as possible'. Exactly what the engineer is doing is a philosophical problem which might lead on to talk of actions and intentions or might lead us back to Plato and the cave, and so on, but it isn't one we are going to solve here in any new or decisive way.

What Héder has caused me to see is that when I say the following on page 16 of TEK: "…strings are entities and entities are strings, so what a string is and what an entity is depends on what is going on at the time." And "…strings are just entities. In the actual analysis, the meaning of terms should emerge from the context without

HARRY COLLINS

[2] Consider not only the strangeness of strings being able to transfer abilities but also the estrangement of phrases like 'this is a photograph of Ludwig Wittgenstein' etc. (TEK:39 ff).

difficulty." I should have been more technical and said that I was taking a certain philosophical position. This is that the problems associated with how we cut up the buzzing, blooming confusion of the world into objects – why we can have tables that are sometimes tree stumps, and chairs that are also sometimes tree stumps, and why we have green and blue instead of grue and bleen – is taken to be resolved by some combination of Goodman's idea of linguistically entrenched projections and Wittgenstein notion of form-of-life.[3] In other words, I start with a commonsense notion of the world given by the essentially sociological notion of form-of-life – which is, I believe, the best we can do when it comes to talking about the way we cut things up into nameable entities. Under this model, for which the icon might be Wittgenstein's discussion of the meaning of 'game', meanings are not given by sharp edged definitions, they are given by the way we generally see things as a result of our socialisation. It is also the case that what part of our socialisation is drawn on varies from context to context. From within that taken-for-granted reality one estranges oneself from particular features of the world in order to ask specific kinds of question about how we cut things up in respect of certain specific entities; you cannot ask all the questions at once.

4. Communication and Inscription

The danger with the commonsense approach I have just described is that it can be used as an excuse for lazy thinking. To avoid it one must be ready to see a problem if it really does affect the main argument. Héder has certainly pointed out something to me that I simply had not noticed. This is that it is hard to draw a dividing line between condition 3 and condition 4 of communication. I would not care very much if there were just a few borderline ambiguities – 'that is life' – or I should say 'that is forms-of-life' – but the problem seems to go deeper. Thus, condition 4 of communication involves a physical change in an entity in order to create the conditions of communication whereas condition 3 involves the enhancement of a string. Héder points out that in my examples condition 4 comprises plugging a new memory chip into a computer while condition 3 comprises sending a string which is enhanced by the addition of a new computer program. But as he says, the new computer program will effect a physical change in the computer's hard disk so both conditions involve physical change. Indeed, he goes on to suggest, though a little less convincingly, that condition 1 and 2 also involve physical changes in the entity that is impacted by the string. From this he argues that there is no real distinction between the first four conditions since all involve changes to the recipient and, at

---

[3]   I discuss the problem at length, drawing on Goodman and Wittgenstein, in Chapter 1 of Collins 1985.

best, it is a matter of degree. He even argues, though much less convincingly, that the process of socialisation that leads to condition 5 is no different to the other conditions because it can be imagined as an elaborate process of string transmissions that lead to a change in the receiving entity (we do not need to worry about this since he goes on to agree that this would not really be socialisation).[4]

5. Ontological fishing

Now, as far as I can understand, Héder has a reason beyond sheer analytical curiosity for wanting to reveal these overlaps between the categories. The reason is that he wants to introduce a thing called a 'pattern processing system' – a computer – as a distinctive ontological feature of the world. He says that if we introduce this third kind of thing we can solve my problem of overlap. We might say he is dangling his favourite ontology in front of me in the hope that I will snap it up as the only way to resolve the problems he has uncovered in my schema:

> If we do not assume the premises that a) P is a pattern processing system, b) the boundaries of the system are clear, c) its embodiment's structure is separable from its dynamically changing state; then we do not have any means to explain the difference between conditions 1-4, because they would differ in only the extent of the change. (Héder 2012:51)

I am not, however, going to take the bait. I have a positive reason for not taking it as I will explain shortly. But, in any case, I am not that hungry. I already have solution to the problem – forms-of-life. I think, however, that I am being a bit more than lazy-minded here because I have spent a long time thinking about this example and have realised something new about TEK as a result. What Héder has made me see is that a set of actors which I thought I had excluded from TEK plays a central role in the book. I say on page 15 of TEK:

> Another unusual feature of the analysis is that very little attention is paid to transmitting entities; nearly all the work of analysis concerns strings and their impact on things with the producers of strings being part of the background.

On page 28 I say:

> In 'telling' the attempt is made to represent lived meaning with the inscribed string. For example, in the case of conversation an attempt is made to represent the meaning

---

[4] I will concentrate on conditions 3 and 4 as that is the most striking example of his point.

as a string comprising vibrations in the air. This book does not deal with the teller or transmitter of a string.

But now I see that the transmitter is already there throughout the book in a latent form and even explicit form as we reach the end of the first half:

> We don't count enabling conditions 4 and 5 of communication as "rendering explicable," because they comprise changes in the receiving entity rather than changes in the string. That is to say, even though a string that initially cannot do work can be made to do work by physical changes in the entity upon which it impacts, we do not say that these changes render the string explicit: *this is just how we use words.* (TEK:81, stress added)

Maybe it would have been better if this had been made clearer. It is the transmitter, drawing on the transmitter's form-of-life, who determines what it is that amounts to a string transmission, what amounts to string enhancement and what amounts to making a physical change to the receiver. That's the general point.

We can, perhaps, narrow it down a bit further. When the transmitter is 'enhancing a string', the intention is to execute one specific communication – the particular multiplication or, to refer to my other example, the particular joke that is being told in the pub. When the transmitter is 'effecting a physical change in the receiver', there is some more widespread increase in capacity is in mind. It may be that the extra inscription on the hard drive increases the computer's capacity to multiply in general but that was not what the transmitter had in mind even though the philosopher can correctly point out that multiplying ability has been improved. The pub joke example is clearer I think – I do not see what general capacity is improved by telling one joke at full length. So it turns out that I am relying on the transmitter's sense of things in spite of the fact that I said in the book that it did not deal with transmitters.

We can put it this way: the analyst can make out that there is overlap between the condition of communication because all involve what can be construed as string transmission (after all strings are entities and entities are strings so a memory chip can be construed as a string), and all involve what can be construed as change in the receiver if we count things like new hard-drive inscriptions as physical change: this is the penalty of my flat ontology. But the actors who do the transmitting do not look at it this way: they know when they are transmitting a particular string, when they are enhancing a particular string and when they are physically changing some entity by adding another entity to it in order to make it possible for strings to act more efficaciously now and in the future.

It is also rather like this: an analyst like Goodman can point out that there is just as much warrant for claiming grass is grue as for claiming grass is green but no-one other than philosophers have any doubt about the matter. The string equivalent of

green is good enough for the first half of TEK where we are trying to work out what we mean by explication. Furthermore, if the form-of-life were to change so that adding a chip to a computer became thought of as part of an act of communication, I don't think it would make any deep difference; explication would still consist of the same set of things though they would not divide up in quite the same way. But the exact way they divide up does not matter so long as the same overall set of acts amounts to explication. In other words, in respect of explication, the way the world cuts up does not matter except in so far as we are trying to create a correct description of the world of the actors – the world we live in.

And this had better be right! Héder says at one point that I make, and need, a sharp distinction between hardware and software – and I now see that such a distinction is there in various passages in the book. But if I really needed that distinction to be an ontological feature of the world I would be in deep trouble. I wrote my first 'computer program' in 1968 or 1969 and at that time to do a multiplication one had to repeat a series of additions into a specific memory location – in other words, just to multiply one had to write a long string. BASIC was a huge advance in my computing life because that long string was now embedded in hardware (or is it pre-loaded software?), meaning I had only to write a short string. So any distinction between hardware and software is not going to last long in an ontological sense and the 'same thing' that was once done with a long string can now be done with a short string plus hardware. But the computer user still knows when they are writing enhanced software and when they are adding, or buying, better hardware – it is just that the boundary keeps shifting. So the ontology cannot matter even though we can still build a classification of kinds of explication which turns on commonsense categories.

6. Social Cartesianism and the ontology of TEK

Where the ontology does matter is when we get to Social Cartesianism. Social Cartesianism is an analyst's category not an actors' category – it is *my* analyst's category. For some reason Héder seems to think that Social Cartesianism is a matter of *belief* – as though I had plucked it out of the air to suit my specific purpose in TEK.

> It is interesting that accepting this dualism is not a logical necessity but rather a matter of belief, the faith of which is a "hostage of fortune" (TEK:144). Collins allows that at some point – although it is not yet imaginable how – someone will be able to construct a machine that is able to use language. It remains unclear whether this would mean that Social Cartesianism is wrong or that machines are also able to access the world of language after all. (Héder 2012:51)

But though the term, 'Social Cartesianism', was introduced only in TEK, the idea that 'the social' is a fundamental constituent of the world and the consequent dualist ontology has been central to much of my work.[5] I was also surprised to find Héder saying that TEK was unclear about the consequences of the invention of a machine that could cope with language in a human-like way.[6]

For me, then, there are language-speaking humans and there is everything else. What positive reason do I have for not embracing the third ontological category that Héder desires to bring in – pattern-processing systems? It is because I want to establish the ontological continuity between string transmission and transformation and physical cause and effect. The bridging device I use is the analogue computer. Analogue computers are just sets of causes and effects.

> Computers, of course, merge into the world of machines in general. Imagine I am driving a backhoe (or JCB). I move a little lever with my fingertips and through a series of analogues (for example, movement of cylinders and flows of hydraulic fluids), a much larger arm and bucket moves. [...] The backhoe is just an analogue computer being used for something other than computing. To summarize, string transformations and mechanical causes and effects are, to speak metaphysically, just two aspects of the same thing. (TEK:50)

I think this continuity is central to understanding analogue strings and, at this point, I cannot imagine abandoning it.

---

[5] For example it is central to the argument of both Artificial Experts and The Shape of Actions (op cit note 1) and it is the very topic of Collins, H. M., (1998) `Socialness and the Undersocialised Conception of Society', Science, Technology and Human Values, 23, 4, 494-516. Oddly in his footnote 3, Mihály cites this later work and agrees that the ideas were already central to my work.

[6] On page 89 of TEK a table of meanings of 'cannot' introduced and the following pages of discussion are all aimed at explaining what is meant by this and related claims. I doubt if anyone has ever been clearer about what they mean by 'cannot' and what the consequences are of a 'cannot' claim being proved wrong. For non readers of TEK, so long as no-one invents a way of making machines that can handle language in the way that humans handle it – something which is 'literally' unforeseeable – then Social Cartesianism holds. As I explain, I am not a prophet and it may be that such a thing will be invented, or introduced to us by aliens from another planet, in the way that faster than light travel after the fashion of Star Trek may one day be within our grasp. If the language equivalent of 'warp speed' comes about then Social Cartesianism will no longer hold and much of what I have written about the relationship of humans and machines would have to be rethought: I have set out the conditions under which my argument would be falsified – in good Popperian fashion. For the time being, however, the argument does not have to be rethought. Uncharacteristically, Mihály has failed to take notice these pages of discussion that deal specifically and at length with something he says is unclear.

7. Other points

Héder suggests that I might change my mind about artificial intelligence if I were forced to consider the success of a computer in the game of 'Jeopardy'. But then he goes on to explain, exactly and correctly, how I would deal with it. The answer is immanent in The Shape of Actions (op cit note 1) and does involve the substitution of polimorphic with mimeomorphic actions, just as he says.

Mihály complains that TEK does not

explain how the first irreducible society materialised. This lack of an origin story in his text raises a number of questions. If the world of language did not exist in pre-human ages, then when did humans create it? Moreover, collectives, subcultures and languages continue to develop today. How many worlds of meaningful languages are there? How many people must speak a language to make it meaningful? Can we observe the jump through the metaphysical gap? (Héder 2012:52)

One small book cannot deal with all of this but there are some snippets in the extended work of myself and others. But, as for how the first languages arose, in a paper entitled 'Building an Antenna for Tacit Knowledge' that is due to introduce a forthcoming special issue of *Philosophia Scientiae* devoted to TEK, I suggest that the best we can do at the moment is imagine it as caused by something like the appearance the mysterious black obelisk visualised in Stanley Kubrick's *2001*. There is a wonderful puzzle there for someone to solve but it is unlikely to be me. How big does a society have to be to support a language? I have inquired of linguisticians and found that their studies show that as societies become smaller their languages become simpler but I do not think anyone has investigated the cut-off point or exactly what would constitute the cut-off point. I have also discussed the matter with my philosopher colleague Martin Kusch and we agree that Wittgenstein's 'private language argument' does not do the job. This is another wonderful topic. How many worlds of meaningful language are there? Among other works, my, 2011 `Language and Practice' discusses the question in terms of the fractal model. As for languages which continue to develop today, *Changing Order* (op cit note 4), investigates that problem by looking at how scientists develop new concepts – which is, of course, a matter of developing new languages.

8. Conclusion

Héder has given a lot of thought to TEK and made me re-think certain things and see them more clearly. I had not noticed that, when thought about in terms of the physical make-up of the devices, condition 3 involved a change in the receiver of a

similar order to condition 4. I did not realise the extent to which I was invoking, and needed to invoke, the sender in order to maintain the difference between conditions 3 and 4. This difference is maintained by invoking actors' categories – or commonsense meanings within a form-of-life: actors' ways of being in the world establish when they are transmitting a string and when they are changing an entity. Though 'context' is invoked in TEK (16) for deciding between the string usage and the entity usage I had not realised that the role of forms-of-life and Wittgensteinian philosophy was a bit more central to my account than I thought it was – Héder's criticisms have brought this out. In contrast one can see the extent to which Social Cartesianism is an analyst's category.

In spite of Héder's criticisms, given this re-thinking, I believe the ontology of TEK is both valuable and correct. There is a flat ontology for strings and entities – they are continuous with one another, just as hardware and software are continuous, which means that, ontologically speaking, string transformation and physical cause and effect are the same. As I say somewhat wistfully:

> …if all the work that is today done with computers was done with elaborate versions of Charles Babbage's Difference Engine, with its clunking gears and ratchets—and it is only logistics that prevents it being so—it would be much easier to understand that a computer is a physical mechanism. (TEK:28-9)

This means that analogue strings and analogue computers can be understood whereas if computers were added as an ontologically distinct category one would, I believe, find all this getting mixed up.

There is, however, a sharp discontinuity between the physical world on the one hand and humans and their languages on the other. Having been forced to think about it, what is going in TEK is that the ontologically mixed-up category of symbols has been rejected for the flat ontology of strings while the hidden ontological function of symbols has been reintroduced in the clear distinction between non-humans and humans. This difference is also expressed in the distinction between physical effect plus strings versus meaning plus language. Héder's discussion, and the contrast with Polanyi that he brings out, intimate, then, that the idea of strings might serve a more important role in the understanding of explicit knowledge than that of merely opening a space for the discussion of tacit knowledge found in the second half of the book. Once more, I thank Héder for this most interesting and provoking discussion of TEK.

References

Collins, Harry. 1985. *Changing Order: Replication and Induction in Scientific Practice*. Beverley Hills & London: Sage.

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

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

Collins, Harry. 1998. Socialness and the Undersocialised Conception of Society. *Science, Technology and Human Values* 23(4):494-516.

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

Collins, Harry. 2011. Language and Practice. *Social Studies of Science* 41(2):271-300.

Dreyfus, Hubert. 1972. *What Computers Can't Do*. New York: Harper and Row.

Héder Mihály. 2012. Explicit Knowledge in the Philosophies of Harry Collins and Michael Polanyi. *Polanyiana* 21(1-2):45-58.

Selinger, Evan; Dreyfus, Hubert; Collins, Harry. 2007. Embodiment and Interactional Expertise. *Studies in History and Philosophy of Science* 38(4):722-40. Case Studies of Expertise and Experience special issue, ed. H. Collins.