EXPLICIT KNOWLEDGE IN THE PHILOSOPHIES OF HARRY COLLINS AND MICHAEL POLANYI

MIHÁLY HÉDER

Department of Philosophy and History of Science
Budapest University of Technology and Economics
mihaly.heder@filozofia.bme.hu

© POLANYIANA 2012/1-2. (21): 45–58.

ABSTRACT

In this article I analyze Harry Collins' explicit knowledge theory. Collins proposed this theory in his book *Tacit and Explicit Knowledge* and applies an undoubtedly novel approach in analyzing the supposedly straightforward concept of explicit knowledge. He discovered that while tacit knowing is natural and ubiquitous, it is explicit knowledge that is strange and needs more scholarly attention. In order to characterize the process of knowledge explication he focused on the physical transfers that take place during communication. His central concept is the *string* that is transferred between the parties involved when communicating. In my critique I bring to light the limits of this approach by proposing examples that are hard to understand based on string transfers yet remain relevant to the original question. I then attempt to define the Polanyian concept of explicit knowledge, which I think can be easily reconstructed based on *Personal Knowledge.* Finally, I call the reader's attention to the major differences between Collins' and Polanyi's theories on tacit and explicit knowledge.

Keywords: tacit knowledge, explicit knowledge, strings, social cartesianism.

1. Introduction

Harry Collins wrote *Tacit and Explicit Knowledge* (TEK) to clarify the title's much-debated notions. His main strategy for reducing ambiguity is to reduce these two notions. He investigates the physical processes of knowledge transfer, because the physical aspect is much easier to observe and discuss. He believes that the mysteries surrounding abstract terms, such as icons, signs and codes, divert the investigator's attention from what is most important as well as easiest to handle: the physical relations between physical entities.

Rather than begin with the usual question of what tacit knowledge is, Collins aptly turns the question on its head and asks what explicit knowledge is (TEK:7). By clarifying explicit knowledge, he can define and handle tacit knowledge more easily. As he points out in the 2011-12/1 edition of *Tradition and Discovery* (Collins 2011), previous literature does not address the notion of explicit knowledge nearly as

often as it discusses tacit knowledge, yet he feels the former is equally interesting if not more so. To remedy this deficiency, he dedicates the first three chapters of TEK to the explicit. Then, the second three grapple with the tacit. The present article will not discuss at length Collins' system of tacit knowledge so it can focus on his concept of explicit knowledge addressed in section 2 of this article.

To give Collins's fundamental assumptions some context, in section 3 I compare them to Michael Polanyi's book *Personal Knowledge* (PK). I will attempt to reconstruct the Polanyian concept of explicit knowledge by mainly examining his "Articulation" chapter. This comparison will help me to arrive at some broader conclusions about communication theory I include in section 4.

2. Explicit Knowledge in Collins' TEK

For Collins, explicit knowledge is knowledge that has been successfully explicated. He aims to describe the process of explication, and he then enumerates what can set back or prohibit this process. By explication he means the demonstration of reduction of the capabilities in question to the transfer of so-called strings (see the next chapter).

The identification of different kinds of obstacles of reduction allows him to set up a classification of tacit (inexplicable) knowledge (TEK:1).

2.1. Strings

To carry out the reduction, Collins needs a firm foundation, so he sketches a universal metaphysics with two central concepts, *strings* and *entities*. Both are physical beings. It is important to note that his string notion has nothing to do with the String Theory of physics or with the popular, versatile data type of computer programming, also called string. This string is a piece of matter that is *inscribed* with *patterns*. It can be anything that is neither featureless nor random (TEK:16). Collins makes no sharp distinction between string and entity, so the strings can be entities and entities can be strings, depending on the context (TEK:16).

Strings interact with entities in different ways. They can interact physically, such as when two billiard balls collide. Collins' system includes a special sub-type of physical interaction, which is when the pattern on a string is *inscribed* to the entity, such as when a printing press embeds letters into paper.

To understand Collins' explicit knowledge notion, one must first examine his concept of communication, which is another sub-type of physical interaction (he does not make entirely clear whether all communication is inscription). He explains, "*A communication takes place when an entity, P, is made to do something or comes to be able to do something that it could not do before as a result of the transfer of a string.*" Sometimes, communication is easy; other times, one needs to exert much more

effort to make it successful (TEK:21). Collins distinguishes five different conditions for successful communication that represent five levels of effort from the easiest to the hardest.

2.2. Conditions of communication

In Collins' enabling *condition 1*, everything is in place on the receiver's (P) side. One single string transfer to P leads to immediate success and enables P to do something it could not do before. This condition probably seems straightforward, but Collins' example is worth discussing here.

In his example, a human enters the formula of an arithmetic addition into a computer that immediately produces the result. In this case, P is the computer. It can perform an action that it never could do before after receiving the signals from the keyboard, and this action is displaying the result of the calculation. We should note that this is exactly what the *made to do something* part in the definition refers to. In this communication definition, the sender can have control over P; moreover, only the sender can evaluate whether the communication was successful or not.

In case of *condition 2*, the simple string transfer does not lead to success. An additional transformation of the string is necessary for success. In Collins' example, a full stop in a printed text might contain a sentence or instruction in micrometer scale. Showing the stop unchanged will not result in success, but with a magnifier, the communication can be made successful. At this point, Collins decides not to specify who is carrying out the transformation.

Moving on to *condition 3*, the transfer of an additional string is necessary to make the original string work. For example, if my computer does not have a calculator program installed on it, then I must install one (additional string transfer) before the original string (the arithmetic formula) will work.

In *condition 4*, even the transfer of a longer, complemented string does not result in success, so a "fixed" physical change in P is necessary. For instance, I might need to add a new memory module before I can install the calculator program. As another example, a weight lifter might be properly instructed how to lift a 120 kg weight, but weight training still needs to take place for successful lifting.

Finally under *condition 5*, further change is needed beyond physical change. To make the string transfer successful, P would have to receive via strings the ability to understand the language. For Collins, this case is different from the other four in a fundamental way. His position is that language ability and other social abilities are different from any other abilities. Later, he describes why he does not believe that language ability can be transferred via strings. He claims that socialization is the only known way that can enable a P entity to understand a language.

2.3. Curious cases of communication

Collins tries to define communication and its conditions in the most comprehensive way, probably because he wants to give these definitions a strong reasoning power

in various situations. If one has reasoning power, then one can use them to evaluate cases of communication that Collins did not mention.

Before we begin, let us note that although "fixed" physical change appears in the description of condition 4, all kinds of physical transfers, including the transfers in condition 1, must result in a physical change in this system of thought. That is, in every case of communication, a string sent to P makes it become P', and the latter entity is able to do something that P could not do. The conditions of the communication designate either the extent of the change or the nature of the change; otherwise, we would be able to determine where P ends and the outside world begins. Why should we accept that installing software on a computer is not under condition 4 when we know that the installation process also results in a permanent physical change on the hard drive of the computer?

If P is really a similar physical being to strings, then we have the following degrees of physical change.

Let us analyze the example of an individual using a calculator to add up numbers. In condition 1, the incoming string (electric signals in this case) is intermingling with the electrons that were already there, which ultimately results in a change on the liquid crystal display (LCD). Then, another impulse (from a different button, like C) clears the LCD and sets the calculator into a similar state as its initial one. (Yet surely, it will not be precisely the same physically.)

Complying to condition 2, the transformation results in a larger physical change in the string, but the effect of this modified string is similar to the effect in case 1. Condition 3 results in some kind of memorization (the additional string must be remembered or understood somehow) that takes the form of fixed physical change. Condition 4 seems to be similar to condition 3, but it appears to manifest to a larger extent. Here, a new version of the previous question arises: should we always assume that the physical modification of P according to condition 4 enacts a successful communication? Or, should we think that inserting a memory module is a successful act of string transfer? If not, how is this case different from installing software?

Collins does not want to give detailed descriptions of the exact characteristics of entities, because he hopes that the difference between the kinds of entities can be established by the investigation of the string transfer question (TEK:15). Furthermore, his baseline theory asserts that anything can be a string or an entity, depending on the context (TEK:16).

Because of this, the distance between P and P' seems to be arbitrary and so do the boundaries of different P-s. Thus, a critical reader might interpret the following as successful communication: a man in a traffic jam is told that he can escape the situation by flying away. This string transfer does not result in a successful communication, but with a certain modification of P – putting a jetpack on his back – communication becomes successful and P' flies away. Yet the question of whether this kind of change is allowed under Collins' fourth condition remains.

If P' was still unable to fly and became only P+jetpack, then would this change be valid? To take another example, let us say that P is an old, gigantic computer, and it must undergo a change in order to respond to users in Hungarian. If I were to hide inside this computer and type answers to the user's messages, then would this be a valid change? Most probably, Collins would reject this example, because P (the computer) was absurdly modified or because my modification was only a trick and not a true modification at all. Therefore, one should be able to establish the difference between inserting a memory module and inserting a human.

Quantitative questions also arise. I can make P, a plastic plate, play chess by soldering integrated circuits and other elements on it and finally installing software. Between P and P', there is a huge gap, but it seems that this process might qualify according to the theory of communication. It would seem even more valid when one considers the case of building and programming a bicycle-riding robot (one of Collins' examples), which seems to be similar. Perhaps if we had a means to disqualify hardware changes but allow software changes, then we could make the distinction.

Another question is whether an organization or a group can be an entity. To my best knowledge, no one in our department speaks Japanese, so we could not read a Japanese letter. However, if the sender of the letter also sends along a Japanese interpreter, then our organization is made to understand Japanese – without socialization. In this case, we can say that a university department is not an entity or that a Japanese interpreter is not a string or both. However, such arguments would be much more restrictive than what is indicated in the phrase, "strings are entities and entities are strings" (TEK:16).

Perhaps this foreign language concept of communication includes cases of ostensive teaching. If I ask a foreigner in downtown Budapest, "Hány óra van?" then the communication will probably not be successful. However, if I also point repeatedly to the part of my arm where people usually wear their wristwatch, combining the previous air strings with visual strings, then the foreigner will be able to show me the time. She might also learn this Hungarian phrase. [1]

Finally, questions might arise about the time dimension of communication. Can someone say that a child's growth from conception to adulthood is nothing other than a long series of string transfers? This definition seems to be formally plausible if we consider an isolated child with only one parent. With an argument like this, someone could make condition 5 transfers look like mere variations of condition 4 transfers, thus attacking Collins' Social Cartesianism (see later).

The counter-arguments to this example could be that 1) only a collective of humans can transfer the ability to use a language and that 2) children are born with the essence of the abilities necessary to understand language, which then only

[1] One wonders whether the scene with the two men on the front cover of *Tacit and Explicit Knowledge* represents this ostensive definition.

needs to be developed. With the second argument, one would also state that the proper language skills of the human as an individual must precede the language of the collectives.

But Collins, by closing down the possibility of condition 5 communication from animals and computers, probably only means that it is impossible for someone to construct a machine or somehow modify an animal until the resulting P' entity will understand language.

In addition, there might be a way to disqualify the transfer of the Japanese interpreter from conditions 3, 4 or 5 by stating that strings must be artificially created by the sender party. Specifically, the sender of the Japanese letter must inscribe patterns on some medium and then transfer it. For example, she could type a formula into a calculator, write software, etc. More precisely, she might just buy the software from someone who artificially created it. After expanding the notion of the artificial, we would probably find that artificial things must be intentionally created by inscription. Sending a Japanese interpreter is just as much a misunderstanding of string transfer as placing a human into a room-sized computer. Human relocation is a natural event and does not involve string inscription. On the other hand, a bicycle-riding robot is entirely artificial, so we might say that only the latter is a result of successful string transfer.

By following this chain of thought, we could say that raising a human child is not string transfer or string transfer accompanied with physical change, because the physical change (in condition 4) must be artificial. The development of a child goes on rather naturally, and its successive stages are not caused by the parents in the same way as the development of a robot is caused by an engineer.

2.4. Hidden ontological level in Collins' definition of communication

It seems that Collins allows (it is difficult to resist using the term "explicitly" here) everything in the category of entity and string. In reality, only humans, computers and possibly animals are implied to be Ps. His examples also feature all three (e.g. TEK:9). Moreover, it appears that he thinks these entities possess a certain structure: they are beings with input channels and an inner state, and they possess the ability to process patterns. In effect, valid entities resemble computers. In *condition 1*, the sequence is input –> processing –> output. In *condition 2*, the input is incompatible, so it is converted first. In *condition 3* the input is not enough, so additional input is needed. Up to and including *condition 3*, P remains unchanged (for a moment, let us set aside my criticism in the previous section about the necessary change of P to P'), because, curiously, people usually do not consider the inner state a physical part of the system when they discuss systems. In contrast, the system itself must be changed in *condition 4*. This change can be the insertion of a memory module, the development of muscles or similar alterations. If we fully accept the hidden premise that Ps are pattern processing systems, then the man with the jetpack is not a problem anymore. He is part of a flying second-order system containing two

individual systems, and one system is controlled by the other. Under this assumption, conditions 1-4 fall into place.

Moreover, the statement that everything can be an entity but also a string is still defendable under this assumption. For example, when programming a pacemaker, the pacemaker is a P entity that is made to do something as a result of the communication. When a surgeon places it in a patient, the pacemaker becomes a string that makes the patient play tennis again. (Alternatively, this can be interpreted as a physical change under condition 4).

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.

2.5. Social Cartesianism

Condition 5 is fundamentally unlike the previous conditions. Collins makes it clear right from the start that he assumes a Cartesian position on this final condition. This position is a kind of dualism in which there are two worlds, the *meaningful world of language* and the *physical world of strings* (TEK:28). Between the two worlds lies a metaphysical discontinuity. Just like in any kind of dualism, the problem of interaction between the two world surfaces here, and Collins addresses it immediately in the "Confounding Strings and Languages" subchapter (TEK:27-9). Here, he explains that the sender of the message makes an effort to transform the language to strings. Then, it is transferred in the physical ways that strings are usually transferred. At the end of the transfer, it is interpreted by the receiver, which – unlike pattern processing – is an operation of more than a physical nature.

Creating a *condition 5* communication would involve the transfer of a non-physical capability of interpretation. Currently, Collins sees no way to do that other than socialization. This metaphysical discontinuity also explains why bringing up a child (socialization) is not merely a sequence of string transfers and physical modifications. Instead, the child is affected in non-physical ways by the community that enables her to access the non-physical world of language.

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.

Collins believes social relations are irreducible, and therefore language is irreducible (TEK:124). He talks about two kinds of beings (although according to my analysis, there is a third kind, the pattern processor system) and the interactions between

these kinds of beings. However, he does not explain how the first irreducible society is materialized. The lack of a story about the origin 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?[2]

2.6. The notion of explicit

The hardest part of reading TEK is accepting Collins' ambiguous ontology. Once we have done so, we can understand the notion of explicit knowledge with relative ease. Explicit knowledge is any knowledge that can be transferred via strings (e.g. *condition 4* communication is not explication, because it changes P. *Condition 5* would probably also change P if this condition had the ability to do so.)

This notion also supports that in Collins' worldview, there is a first, hidden gap[3] between pattern processor systems and other physical things. Otherwise, the pattern that is stored inside a system would require an immediate a physical change. In *chapter 5*, we can see that he makes a strong distinction between hardware and software (TEK:100). In this sense, the construction of the bicycle-riding robot cannot be explication. However, if we already have a robot with the proper structure, programming it *is* explication. This also means that the knowledge of riding on a bicycle is assumed to be in the software and not in the software-hardware unit; otherwise, it would not qualify as fully explicable, because creating hardware is not explication.

In my opinion, this part is the second most problematic assumption in *TEK* after Social Cartesianism. On the one hand, there is no metaphysical discontinuity between strings and entities as long as entities are not humans – on this end, we have a flat ontology (See Lowney 2011). On the other hand, the non-human P entities can be made to do things in the physical world after receiving physical strings, but P in this case remains physically unchanged, because it is a pattern processor system that interacts with the physical world with non-physical properties, such as inner states.

2.7. The evaluation of AI

An interesting analogy in the book is Collins' comparison of the human's capacity to use language to one of the simplest computers in existence – the Chinese room. Referring to this computer doesn't seem to be an effective argument, because the

[2] Some answers to these questions might be reconstructed from Collins' *Socialness and the Undersocialized Conception of Society* (1998), in which many central ideas of TEK is already present.

[3] A second gap lies between humans and all other phenomena.

Chinese room is stateless, has no recursion and uses no feedback. The combination of these qualities with this computer's unconditional nature and single input channel makes it unlikely to produce anything interesting.

The definition of a Chinese room is a machine that can only answer to pre-programmed questions. Its role in Collins' book might be to represent all kinds of computers while still being easy to understand. However, the Chinese room does not serve as an analogue for the majority of real computers that have qualitatively more computational power as the program languages for real computers are usually Turing-complete.

It would be interesting to evaluate a more successful system in Collins' framework, for example the IBM Watson project. This system was specifically designed to answer questions it had never seen before. It does this as a part of a general knowledge quiz game, titled Jeopardy. The questions are enriched with language tricks and jokes. It is well known that Watson won the Jeopardy championship in 2011 against former human champions (TEK was published in 2010, so unfortunately the analysis of this case could not be included) The game was played in English, so one might ask whether this means that the capability of understanding language was transferred to a machine. An even more interesting scenario is a similar situation of the "beer-mat" knowledge (TEK:59). This case is about a man who reads a short text about holograms from the back of a beer mat. It does not allow him to build a hologram, but he could answer questions about holograms, such as, "Do holograms include lasers?"

Watson was trained on mostly Wikipedia pages that are similar to the back of a beer-mat. Yet, not all the possible questions of Jeopardy were programmed into it, and it still won. Does this mean that Watson reproduced the man with the beer mat and the language skill that comes from socialization? In other words, was the act of programming Watson some kind of limited socialization?

Collins' answer would probably be "no". Maybe he would argue that this is a typical case of a so-called polimorphic action that turned out to be reproducible by mimeomorphic actions (read: can be carried out by a Chinese room), just like playing chess. This way, playing Jeopardy became explicated, but it is not the same as the knowledge of using language. However, this also means that this game cannot be seen as a real social situation anymore. In Lowney's opinion, Collins' effort helps us to remember what makes us different from machines. I am afraid of the opposite possibility: as the front line of explication progresses in this system, the set of things that make us unique is continuously shrinking. We can only hope or expect (as Collins does) that one day, this front line will simply stop, but as long as the expansion of the explicit continues, any human activities that are reproduced will be devaluated.

Anyhow, it is undeniable that Collins brought a brand new approach into the debate over tacit and explicit, and this approach came with its own metaphysics. *TEK* is a novel project that aims to solve a number of slippery questions, including the explicit, the tacit knowledge and the evaluation of AI, in a unique way.

3. Explicit Knowledge in Polanyi's Philosophy

Collins' complaint that explicit knowledge does not have the backing of proper literature and a straightforward definition is by no means unfounded. For instance, in *Personal Knowledge*, Polanyi simply starts using the term "explicit" in a footnote yet does not include later in the book any definitions of this term. Meanwhile, Polanyi asserts that wholly explicit knowledge is "unthinkable". In order to understand this assertion, a proper definition would indeed be helpful.

I believe Polanyi's chapter titled "Articulation" offers the most important information about Polanyi's concept of explicit knowledge. From the introduction of this chapter, we learn that Polanyi feels it is evident that there is a huge gap between the intellectual capabilities of humans and animals. The use of language that is enabled by articulation creates this gap. Collins and Polanyi would agree on these overarching points. However, there are huge differences in the details of their theories. While Collins believes the most important element of language is the collective, Polanyi defines language as being the result of the articulation capability of the person, which also has a social aspect.

Polanyi explains the gap between animals and humans by telling the case of researchers who observed the parallel development of a chimpanzee baby and a human baby. In the initial stages of the subjects' lives, the researchers observed the human baby having no significant advantage over the chimpanzee baby. However, the human child's intellectual power apparently multiplied when he started using language, which Polanyi believes gave the child access to the cultural heritage of his ancestors (Polanyi 1962:70). In response to this study, Polanyi insists that the biological facilities that enable language are not dramatically different from those of the chimpanzee, so he sees no perplexing jump in evolution between the two species. Instead, he thinks that certain inarticulate facilities are already present in animals and that only the combined capability of these facilities is missing from the animals. The result of the right combination – the capability of articulation that is necessary for developing language skill – is what makes humans intellectually superior.

3.1. The degrees of pre-lingual intellect

Polanyi differentiates various intellectual capabilities by relating them to different concepts of learning, which he dubs Types A, B and C.

An example of type A learning is a rat that learns by chance how to obtain food by pressing a lever in a laboratory experiment. Given its desire for food, it will repeatedly press the lever after learning the effect of doing so. Polanyi calls this trick-learning, whereby the animal learns how to perform a particular operation.

Type B learning features signs that enable the animals to predict a certain future event or state. A rat is capable of learning that the boxes with a certain sign contain food. The instant the animal discovers the relevance of the sign, it no longer

bothers to open the boxes without the sign. In addition, Pavlov's dogs are capable of predicting food by noticing a sign (of course, the experiment is usually cited to show how unconditional a reaction can be).

Type C learning is latent learning. A rat is able to learn the shortest route in a maze after exploring it for some time. Even when the experimenters place an obstacle in this optimal route, the rat finds a good alternative route much more quickly than a trial-and-error method would allow. This means that the rat must have a mental representation of the maze that it can apply to an altered situation. The aim of these examples is to show the different stages in the development of representation.

3.2. Language

Humans are capable of even more than the latent learning in the previous example. The source of that ability, according to Polanyi, is the *certain way* humans create and manage mental representation. He calls the principles by which these processes work the Operational Principles of Language, because these principles result in the appearance of language when they function properly.

Humans are capable of effectively mapping the world to a relatively small set of elements. This capability is called the Law of Poverty, and Polanyi calls these elements *symbols*. However, a small number of elements is not sufficient for communication. Their mapping to the experience about the real world must also be *consistent*. Moreover, the combination of these symbols should not be arbitrary, because a certain *grammar* is needed.

The manipulation of the symbols should be rather effortless so that it falls under the Law of Manageability. This law must be in effect from the *primary denotation*, the *reorganization* and in the *reading of the result.*

Polanyi's theory holds that if symbols meet these conditions, then they will enable the skill of articulation, which, in this book, means the process of creating symbolic representation (Polanyi 1962:82-5). The skill of articulation and the skill of managing symbols are simply further developed versions of the animals' similar skills.

In his system, skill is one type of knowledge, and the potential for each skill is embedded in the human brain. As Polanyi puts it, tacit skills are cooperating with the explicit (Polanyi 1962:87). This statement makes clear what Polanyi considers *explicit*: the symbolic representation that is part of a language and that was created by articulation. Knowledge that cannot be articulated in terms of the symbols of a language is tacit. Not everything that is articulated meets the operational principles, so articulated and explicated are not always the same in Polanyi's framework.

If we are interested in assessing only the concept of explicit, it is not necessary to accept Polanyi's account of the emergence of language. Without that account, we can still summarize what he calls explicit: everything that is expressed in linguistic symbols. *Language* is inclusive here: it can be text, formalisms, characters, diagrams, charts, etc. (Polanyi 1962:78). Nor it is necessary to accept the operational principles, according to which pyramids are symbols but not parts of language because of the

pyramids' lack of manageability (Polanyi 1962:81). In this chapter, I will use the term *explication* in this sense, which contrasts strongly with Collins' concept of explicit.

According to Polanyi's theory, a person can never fully express her whole knowledge with symbols, because even her most formalized knowledge is still partly made of the facility of articulation as well as of consistent denotation of those symbols and the facility of managing them. These are not symbols, so they are not explicated. Yet an external observer (Polanyi's neurobiologist for instance – see later) can assign symbols to these facilities of the brain, and those symbols can be the explicit part of the observer's knowledge. But, this is knowledge *about* someone else's knowledge, not the external observer's own knowledge.

This is why Polanyi believes it is trivial that no one can explicate her own bicycle riding skill in human language. Even the question of the possibility of fully explicit knowledge only surfaces with the strongly formalized, propositional sentences, such as knowledge of Laplace's demon (Polanyi 1962:139-41) or logical sentences that are processed by computers. According to Polanyi, even these are not fully explicated as the symbols cannot include their own meaning that relies on the tacit facilities of consistent symbol denotation, etc. Even for understanding the most formalized sentences, a person is needed. As Laplace's demon has no personality it does not have real knowledge of the world, even though it explicitly knows the position and momentum of all particles of the universe.

By emphasizing the impossibility of knowledge without a person, Polanyi tries to call attention to how hopeless and even dangerous are those programmes that promote the ideal of the fully objective knowledge, like neo-positivism.

It is important to note that when Polanyi talks about inexplicability, he consistently means from the perspective of the knower. He sharply distinguishes inexplicability from another situation in which a scientist examines the processes unfolding within a subject while the subject is performing something (e.g., riding a bike) with his knowledge (*Polanyi* 1968:39). The scientist can see and describe those inner processes that are hidden from the subject's attention. This is explication, but not of the knowledge of the subject. This is an explication of the knowledge of the scientist (maybe the scientist cannot even ride a bike). Furthermore, this explication is only partial – the scientist cannot fully explicate everything he knows about the bicycle rider's skill.

The inexplicability of riding a bicycle in this framework only means that no one can learn to ride a bicycle by simply reading books about the activity. Performing the act is only possible by exercising balance, which builds the necessary facilities in the performer's neural system that cannot be made by receiving language sentences[4]. Now, there are other imaginable ways of riding a bicycle. For instance – as in Collins' example – one might learn to ride a bicycle on a low-gravity asteroid by

[4] For Polanyi, these are different kinds of changes, because the body and the symbols are on different ontological levels in his emergent ontology.

simply reading and interpreting some rules regarding the angle of the bike, the handlebars and so on. Learning about this type of activity would replace a certain kind of tacit-only knowledge with a more developed form of explicit+tacit knowledge that involves reading and interpreting, that is, inexplicable abilities that are present in only humans. Learning to ride on an asteroid this way would not engender full explication, as it involves tacit knowledge about reading and following instructions.

Polanyi developed this conceptual framework to show that the knowledge of the bicycle rider and the knowledge of the scientist share traits, no matter how formalized and symbolic the latter's knowledge is. That is, neither of them can eliminate their own person from the knowledge. At the same time, knowledge can be replicated in his system even though it cannot be explicated. One would replicate it through ostensive learning in the training of medics, riding exercise bikes, etc. Moreover, animals have knowledge, but they have no language, so everything they know is tacit[5].

Without going into further detail here, Polanyi explains that acquiring knowledge is always an act of belief. It cannot be a result of logical inference, because knowledge cannot be fully formalized in logical sentences in the first place. Therefore, the proposer of the knowledge must exhibit persuasive passion to compel the listener to jump the gap that cannot be bridged with logic alone.

Meanwhile, knowledge has truth value in Polanyi's philosophy, regardless of its inexplicability. This truth value is neither subjective nor objective. With the sub-title *Towards a Post-Critical Philosophy*, he wants to express that he is against objectivism as well as the relativism that criticizes objectivism. He believes both approaches are wrong and can even be dangerous[6]; instead of following those, he suggests following a third way that is now called the personalist or emergentist approach.[7]

4. Conclusions

As we can see, the notion of explicit and tacit largely differs between the philosophies of Collins and Polanyi. These authors offer independent, incompatible answers to more-or-less the same questions.

In Collins' system, the successful replication of knowledge by transferring strings means that knowledge has been explicated. At the same time, his position is that the knowledge that involves understanding language or other social relations cannot be explicated via string transfer. However, if there exists a string that makes something ride a bike, then bicycle riding would be explicated.

[5] In our previous paper, which we wrote with Daniel Paksi to the *Appraisal Journal (Héder and Paksi 2012)*, we also argue that certain machines are similar to animals *in this respect* (not in others).

[6] See his moral inversion concept (Polanyi 1962:233).

[7] It would be interesting to discover in detail the parallels between Polanyi's program and the program of the Third Wave of Science Studies.

In Polanyi's system, explication means expressing something in language, even though one cannot articulate any kind of knowledge fully. He believes communicating complete knowledge is completely impossible because every kind of knowledge is, in part, embodied in the person and that embodiment is metaphysically different from the signs of language and therefore cannot be articulated. Knowledge can be replicated, but replication is not explication.

We can conclude that Collins built a completely novel conceptual framework for tacit and explicit knowledge. One must accept Social Cartesianism to use Collins' theory[8]; otherwise, there would be no inexplicable knowledge, not even language skills. Knowledge that does not involve language can be explicated, and it is only a matter of time, effort, interests and logistics before they are explicated. In contrast, Polanyi simplifies tacit and explicit knowledge. Everything that is represented in language is explicit, but language always conveys less than what we know. Therefore, explication in Polanyi's theory is always incomplete.

Acknowledgement

I am very thankful for the comments of János Liska, Daniel Paksi and Margit Ronkay. This work was supported by the grant TÁMOP - 4.2.2.B-10/1--2010-0009.

References

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

Collins, Harry. 2011. Analysing Tacit Knowledge: Response to Henry and Lowney. *Tradition and Discovery* 38(1):18-37.

Lowney, Charles. 2011. Ineffable, Tacit, Explicable and Explicit: Qualifying Knowledge in the Age of "Intelligent" Machines. *Tradition and Discovery 38(1):38-42.*

Polanyi, Michael. 1962. *Personal Knowledge*. London: Routledge and Kegan Paul.

Polanyi, Michael. 1968. Logic and Psychology. *American Psychologist* 23:27-43.

Héder, Mihály and Daniel Paksi. 2012. Autonomous Robots and Tacit Knowledge. *Appraisal* 9(2):8-14.

[8] This fact is openly admitted by Collins in TEK:164-5.