



A Philosophy for Our Time?
A Comment on Paksi and Héder's *Guide to Personal Knowledge:*
The Philosophy of Michael Polanyi, Tacit Knowledge, Emergence and the
Fiduciary Program

Alessio Tartaro

On Guides to Philosophical Texts

The guide to the philosophical text is itself a philosophical genre. Commentaries on major philosophical texts already existed in ancient Greece (see, for example, Alexander of Aphrodisias's commentaries on Aristotle), and even today guides to major works in the history of philosophy are an essential tool in philosophy courses. But why write a guide to a philosophical text? What is the relationship between the text and the guide?

Philosophical guides are written for several reasons. First, philosophical texts are, more often than not, difficult to understand, even for people who specialize in philosophy. The reasons for these difficulties can be many. Philosophers often use technical terms or use words with meanings different from the commonly accepted ones, which can mislead readers. In addition, philosophical texts have a deep argumentative structure that runs across the entire text. This deep structure can be difficult to grasp for the reader, who ultimately fails to see how the unity of the work goes beyond the sum of its parts. Finally, more prosaically, the philosophical text is difficult because it deals with difficult subjects that are not easily handled by our cognitive, linguistic, and conceptual tools.

Another reason some philosophical works are difficult to fully understand is that the philosophical text invokes a broader philosophical context. Within a text, an author often establishes a continuous dialogue with contemporary and non-contemporary philosophical ideas and positions in order to criticize, support, and revise them. Philosophical texts are frequently filled with these more or less explicit references to the surrounding environment. The reader who is unfamiliar with this environment has difficulty grasping these references, the understanding of which is essential to better understand the philosophical text, and a guide can help promote this broader understanding.

A guide is thus, first and foremost, a remedy to the difficulty of some philosophical texts. Its basic purpose is to provide a more accessible way to follow in depth a text and make a difficult reading easier. A guide stands to the text as a map stands to a region. The map provides itineraries for moving within the same region. By consulting a map, we are able to go from point A to point B without having to wander over that region until we find our destination. This is possible because the map does not coincide with the region point for point and does not reproduce it completely accurately. Like the Map of the Empire in Borges's story *On Exactitude in Science*, a map on the same scale mile to a mile is completely useless, because finding the path on such a map is the same, and as equally complex, as finding it in the region itself.

The same happens in the case of a guide, which, for this reason, cannot fully retrace the original text because otherwise it loses its function. At the same time, the guide does not replace reading the text but

supports it, just as consulting a map does not replace the experience of going on the marked paths of the map. Rather, the map helps us move along those paths. Similarly, the guide should provide the reader with a tool to engage with the original text rather than indulge in elaborate discussions (pro or con) with points developed in the text. To succeed, the guide must present a good balance between fidelity to the original text and the need to say only the important things. On the one hand, a guide cannot overly approximate or mimic the text, because otherwise it will be no easier to understand than the text itself. On the other hand, a guide must not overly simplify either, because this will not truly facilitate understanding. A good guide, in short, must find the right scale.

Finding the right scale necessarily involves selecting relevant aspects and neglecting others. A mountain map, for example, should be able to tell us at what point we will have to take a certain path to reach the summit, but it does not need to tell us that such a path is in a pine forest. Obviously, the author of the map knows that the trail is in a pine forest but considers this information irrelevant to the user of the map. The same happens in the case of a philosophical guide. The author has a thorough knowledge of the text, knows all aspects of it, but must judge what is relevant and what is not in order to provide the reader with an understanding of the text. This process of selection and exclusion is the most difficult task in producing a philosophical guide.

The complexity of a philosophical text and the desire to make it more understandable are good reasons to write a guide. Yet they alone are not enough. In fact, the history of philosophy is full of texts that for various reasons are difficult to understand but for which no one has written guides. If there are guides to Aristotle's *Metaphysics*, Kant's *Critique of Pure Reason*, and Wittgenstein's *Tractatus*, it is not merely because these texts are difficult and therefore need to be supported by a guide in reading them. Rather, the guides exist because these texts are so important in the history of philosophy that it is necessary to provide pedagogical tools to support the reader in personally engaging with them. The creation of a guide to a philosophical text, therefore, is supported by the idea that the original text is of significant importance and deserves to be read, and, since reading it is difficult, the guide aims to facilitate the accessibility of the text by making it more understandable.

On Guide to Personal Knowledge

Guide to Personal Knowledge: The Philosophy of Michael Polanyi: Tacit Knowledge, Emergence and the Fiduciary Program by Dániel Paksi and Mihály Héder (henceforth *GPK*) admirably corresponds to the above ideas about the function, motivations, structure, and limitations of a guide to a philosophical text. The authors offer a faithful and never simplistic or complex account of Polanyi's work, focusing as much on the particular topics as on the general purposes of the book, highlighting its main extratextual references, and thus providing an essential tool to support the reader in discovering the depths of this profound and complex work.

Personal Knowledge is, first and foremost, a difficult book in which Polanyi sets out his ideas on a wide range of topics. The book articulates ideas about metaphysics, philosophy and sociology of science, epistemology, philosophy of technology, philosophy of language, philosophy of mathematics, philosophy of probability, philosophy of mind, political theory, philosophy of biology, and even theology. As the authors of *GPK* rightly state, *Personal Knowledge* "presents a unique worldview" (xv). The vastness of the topics covered is the initial element of difficulty when trying to read *Personal Knowledge*. How do such disparate and diverse topics fit together? What is the common thread that binds them? This is a question that any

reader finds himself or herself asking at some point in reading the book. Paksi and Héder have presented an answer to this question at the beginning of their guide.

Personal Knowledge is “an enquiry into the nature and justification of scientific knowledge” (*PK*, I; NB: citations use 1998 e-book pagination), but it is not only this because, around the new conception of knowledge, Polanyi constructs “a system of correlative beliefs” (*PK*, II). Beginning with a critique of objectivist and positivist views of scientific knowledge, Polanyi constructs an alternative that also has profound consequences for a wide range of questions outside science, ranging from socio-political matters to the idea of evolution. As their answer to the common thread that binds Polanyi’s topics, Paksi and Héder summarize his general goal as follows:

Polanyi’s goal in writing *Personal Knowledge* was precisely to answer this twentieth-century trap by providing a concept of knowledge that enables modern persons to develop acceptable forms of relationship with older inherited traditions, and, at the level of the individual, self-acceptance, which includes a harmonic relationship to our human possibilities. Polanyi aims to help modern persons become at home in our universe.

Part of the effort of *Personal Knowledge* is to create and make acceptable a new idea of the human that is entirely consistent with the concept of evolution. (*GPK*, xvi)

Polanyi develops this goal by resorting to a number of new concepts that the authors carefully present in advance in their “Preface.” These innovative concepts that are treated in the preliminary analysis include “objectivity,” “trust” and “fiduciary program,” “tacit knowledge,” “intellectual passions,” “deceptive substitution,” “moral inversion,” “operational principles,” “logic of achievement,” and “emergence.”

“Objectivity” is crucial to understanding the *pars destruens* of Polanyi’s reflection, in which he criticizes the idea of detached, universal, infallible knowledge and the scientific, ethical, and political consequences of this idea. The term “fiduciary program,” on the other hand, represents the *pars construens* of the work and is fundamental to understanding Polanyi’s ideas about belief, the relationship between knowledge and belief, and the foundation of knowledge in man and society. Against the Objectivist ideal, Polanyi re-proposes St. Augustine’s “*nisi credideritis, non intelligitis*” and the need to rely on beliefs in order to gain knowledge and understanding.

“Tacit knowledge” is undoubtedly Polanyi’s most famous concept, and it has been widely adopted and reworked, even outside of philosophy. In *GPK*, the authors have the merit of clearly reconstructing Polanyi’s original concept, going beyond the famous tautological (and therefore not always useful for understanding) statement, “we can know more than we can tell.” Tacit knowledge describes a form of knowing akin to skill, which has an evolutionary origin, a fiduciary foundation, and broad socio-cultural consequences. Highlighting these dimensions of the concept is crucial to avoid misunderstanding and simplification. Correctly, the authors highlight that tacit knowledge and personal knowledge are not the same thing but that “tacit knowledge is part of personal knowledge” (*GPK*, xix). Another element present in all acts of personal knowing is “intellectual passions,” a sign of the subject’s personal participation in the act of knowing and a fundamental element in scientific practice.

The authors then introduce two concepts fundamental to understanding Polanyi’s socio-political reflection: “deceptive substitutions” and “moral inversion.” According to Polanyi, “deceptive substitutions” are false reasons for why some scientific theories are developed and accepted; these substitutions are meant to conceal the real motivation, namely the satisfaction of intellectual passions. Thus, for example, instead of

saying that one accepts a theory because it is “true,” one says that one accepts it because it is “simple,” thus denying that truth is one of the main intellectual passions and that the satisfaction of this passion, i.e., having true theories, is one of the goals of science. Related to the concept of “deceptive substitutions” is the process of “moral inversion,” which occurs when material purposes replace moral aims in human action due to a strictly mechanistic conception of man and society proper to modernity. According to Polanyi, moral inversion underlies the decline of liberalism and the rise of totalitarianism and fascism in the twentieth century.

Finally, the last two concepts refer to Polanyi’s reflections on machines, living beings, and evolution. The authors effectively summarize these two concepts (“Machines and living beings follow *operational principles* that are oriented toward *achieving goals*,” *GPK*, xxi), highlighting how they cannot be reduced to physico-chemical principles and are fundamental to explaining the functioning and success of these entities. Finally, the concept of “emergence” explains how “these entities come into existence” (*GPK*, xxii). More generally, as the authors explain, “the concept of emergence describes the processes by which higher levels come into existence from lower ones” (*GPK*, xxii), giving rise to a world with a layered structure.

As mentioned above, every guide needs to select and exclude, from the richness and abundance of the text, the elements most appropriate to facilitate understanding. In this regard, the choices of the authors could not have been better. The concepts presented preliminarily form the basis of Polanyi’s reflections and are fundamental to understanding everything else in the book. In fact, so many other concepts that have special resonance in *Personal Knowledge* (e.g., “commitment,” “indwelling,” “tradition,” “conviviality,” “anthropogenesis,” “ultrabiology,” etc.) cannot be explained without resorting to the concepts presented in the authors’ “Preface.” These and many other concepts are explored in depth in the respective chapters.

The book is divided into thirteen chapters that follow the structure of *Personal Knowledge*. Each chapter is introduced by a set of objectives that Polanyi sets out to achieve. The objectives summarize Polanyi’s theses, while the body of the chapter reconstructs the arguments Polanyi uses to support these theses. The arguments are presented in the same order as developed by Polanyi but are introduced by a paragraph title that helps to follow the scansion and development of these arguments. In addition, the text of *GPK* is interspersed with many direct quotations from *Personal Knowledge*. These quotations link the guide to the original text and help the reader become familiar with Polanyi’s terminology and language. The use of figures and tables helps illumine even the most complex concepts, such as tacit integration and the difference between random and ordered systems. In some cases, the authors make use of intra-textual links to connect different aspects of Polanyi’s thinking and show its unity. For example, Polanyi’s critique of neo-Darwinism, initially developed in chapter 3 on the basis that it incorrectly explains the comprehensive order of life with the concept of randomness, is reconnected to chapter 13, where neo-Darwinism is rejected in favor of an emergentist conception of evolution. Similarly, in chapter 6, “Intellectual Passions,” the authors link the concept of intellectual passions to the critique of the Laplacian ideal of objective knowledge (in chapter 1) and the discussion of moral passions (in chapter 7). Through these intra-textual connections, the authors outline pathways within the *PK* text to help readers understand specific aspects of Polanyi’s thought. In conclusion, in reconstructing the content of each chapter, the authors stick steadfastly to the *PK* text and resist the common pitfall of philosophical commentary: digressions that distract the reader from the thought of the author being commented on. In attending to Polanyi’s thoughts, the authors do not wander, nor do they merely paraphrase, but they manage to clearly explain Polanyi’s main ideas.

GPK is enriched by twenty-five boxes that provide information on ideas, authors, and concepts directly or indirectly referenced in Polanyi's discussions in the various chapters. These boxes are a valuable tool because they do justice to the open character of the philosophical text. Frequently, Polanyi refers to contemporary events and theories but without delving into them. For example, in his critique of the objectivist view of science, Polanyi dismisses the role of the Michelson-Morley experiments in Einstein's development of the theory of relativity. In *Personal Knowledge*, however, Polanyi does not extensively develop this argument, and he never explains these experiments in detail. The authors of *GPK* fill this gap with a box (Box 4, *GPK*, 11–12) in which the experiment and its significance are analyzed. In other cases, a box serves the function of developing a comparison between Polanyi and related philosophical figures, for example, Thomas Kuhn (Box 13) and Samuel Alexander (Box 23). By providing context and background for many of Polanyi's reflections, the boxes in *GPK* help a reader understand *Personal Knowledge* more fully and make this guide an even more valuable tool.

To help the *PK* reader develop a deeper appreciation of some of the themes of Polanyi's thinking, *GPK* authors could have provided more bibliographical references to both Polanyi's other works and other scholars' discussions of Polanyi's ideas. This is a recommendation to consider in a possible future revised edition. There is a rich secondary literature that analyzes specific aspects of Polanyi's philosophy. Although an extended treatment of this literature is beyond the scope of a guide, a somewhat richer bibliographical apparatus could certainly be helpful for the reader who wants to discover more about Polanyi.

There is a final issue to be considered: Was a *PK* guide needed? This question is intimately tied to another question: Is *Personal Knowledge* an important text that is worth reading? In the "Foreword" to *GPK*, C. P. Goodman claims that Polanyi's philosophy is "a philosophy for our time" (x). If this is true, then there is no doubt that the publication of *GPK* is an important contribution that brings readers and scholars closer to Polanyi's thought.

In the remainder of this commentary, I highlight a contemporary issue on which Polanyi's philosophy undoubtedly proves to be a philosophy of our time: the nature and limits of artificial intelligence.

Recently, a Google engineer, Blake Lemoine, claimed that the Large Language Model he was working on, based on artificial intelligence, was conscious. The news caused a stir in the media but also among academics and researchers. It provoked many reactions and rekindled a now-recurring debate about the nature of artificial intelligence and its limitations. Among the various contributions occasioned by the news, Jacob Browning and Yann LeCun wrote an article in which they downplay Blake Lemoine's claim and dismiss the idea of conscious artificial intelligence. In the article, titled "*AI And The Limits Of Language*," the authors go further by actually asserting that artificial intelligence, particularly LLMs, not only cannot be conscious but do not even approximate human understanding.

Taken alone, this view suggests that a Polanyian position is alive in the contemporary debate. Witnessing the emergence of intelligence as a field of research in the 1950s, in fact, Polanyi was a bold opponent of the computer-mind analogy, according to which "intelligent behavior is based on a machinery which, in organisms possessing a nervous system, operates on the principles of digital computers" (*PK*, 358). As Paksi and Héder explain, "according to Polanyi, the fundamental failure of the computer-mind analogy is that computers are merely logical inference machines. They do nothing other than transform the programmed, formally symbolized explicit sentences by strict formal rules. [...] in their cases, there is neither assertion nor meaning, only the logical transformation of explicit sentences which in themselves without personal commitments are meaningless" (*GPK*, 113).

Obviously, the kind of “logical inference machine” known to Polanyi was not equipped with the same level of autonomy, interactivity, and adaptivity as modern artificial intelligence systems. However, this does not change the substance of Polanyi’s argument. For example, Polanyi recognizes that adaptive and self-regulating capacities with respect to external situations are functions that can be explained according to a machine-like conception:

The machine-like conception of living beings can be extended to account in principle for their adaptive capacities. An automatically piloted airplane approximates the skills of an air pilot. Its mechanical selfregulation co-ordinates its activities in the service of a steady purpose, and it may even appear to show a measure of resourcefulness in responding to ever new, not exactly foreseeable situations. (*PK*, 353)

In a provocative passage, Polanyi even goes so far as to concede that “It might not be inconceivable that a machine of sufficient complexity would develop conscious thinking, without losing its machine-like character” (*PK*, 353).

However, this does not imply that even the activities of human intelligence can be explained according to this mechanical conception of all vital adaptive functions. Indeed, while it is true that an “organism sustains itself by functioning as a machine” (*PK*, 424), living beings are not reducible to these machine-like operational principles, let alone to their physical and chemical characteristics. Against this idea, Polanyi suggests the presence of an active center “operating unspecifiably in all animals” that is in addition to the machine-like functions characterized in terms of operational principles: “There are then two principles at work in animals: namely, (1) the use of machine-like contrivances and (2) the inventive powers of animal life” (*PK*, 354). This active center is not superimposed or parallel to machine-like functions but is superordinate to it: “while the animal’s machinery embodies fixed operational principles, this machinery would be impelled, guided and readapted by the animal’s unspecifiable inventive urge” (*PK*, 354). For this reason, Polanyi considers living beings as “instances of morphological types and of operational principles subordinated to a center of individuality” (*PK*, 405). Living beings are thus a combination of types, operational principles, and individuality. The active center is an expression of individuality and of a “personhood” present, at different levels, in all living beings (*PK*, 409). An expression of this inventiveness of active centers is, for example, equipotentiality, by which an organism offers a series of solutions for the same technical problems (e.g., mutilated rats able to activate different motor patterns to achieve the same goal) (*PK*, 355). Some aspects of embryonic development also respond to this principle (*PK*, 356). The very emergence of new operating principles in living things depends on the activity of an “orderly innovating principle” (*PK*, 418) whose actions are different “from the conditions which release and sustain its actions” (*PK*, 404). Evolution as a whole is “a process of fundamental innovations, tending to produce ever higher biotic achievements” (*PK*, 404).

It is the presence of this active center, an expression of individuality, creativity, and personality guiding and readjusting the machinery of living beings, that distinguishes them from machines. Should a machine develop conscious thinking, this would still make it different from a living being in that the machine’s “consciousness” could exert no influence on its own mechanical operations: “conscious thoughts would be the mere accompaniment of automatic operations” (*PK*, 353). This idea, representing a form of occasionalism, would be equivalent to saying that Shakespeare’s conscious thoughts have no influence on the writing of his plays, an act that in itself can be explained as the result of mechanical operations. Although this is

“not strictly inconceivable...nobody can believe in it in practice” (*PK*, 354). Shakespeare’s plays are not the product, nor are they explicable as the result of mechanical operations. They are, on the contrary, “a massive demonstration of a creativity which cannot be explained in terms of an automatic mechanism” (*PK*, 354).

The idea of living beings as a combination of (morphological) types, operational principles, and individuality allows us to distinguish living beings and machines, which are instead a combination of types and operational principles without an active principle. One can grant that machines themselves have an emergent structure and that machines may possess knowledge, even tacit knowledge, but this does not bridge the gap between a living being whose active center is a product of innovative evolution and a machine that ultimately relies on mechanical operations and is not animated by any active center. The way a modern AI system interacts with the environment and adapts does not reflect the presence of an active center. Again, it would be possible to attribute its functions to an active center, but although this is strictly conceivable, at present it seems difficult to accept.

One consequence of this reasoning of Polanyi’s is the following: since understanding is a biotic achievement (*PK*, 365), and biotic achievements are those of an active center (*PK*, 425), it follows that no entity without this active center can be capable of understanding. This brings us back to the article by Browning and LeCun and the idea that LLMs, although able to master language like humans, possess no understanding or have only a “shallow understanding.” Systems like GPT-3 are efficient in guessing which next word is most likely and in “coming up with a plausible sentence given the prior line.” In this way, these systems can achieve amazing results such as conducting human-like conversation, explaining difficult concepts, rephrasing, and retelling or summarizing stories. And yet, as the authors state, “the capacity to rattle off linguistic knowledge” is different from “skillful know-how for how to do things like being empathetic or handling a difficult issue sensitively.” While the former can be incorporated, the latter cannot. This is because the second type of skill is not linguistic and thus cannot be introduced into the system through training on words and sentences. Realizing how much of human knowledge is not linguistic is the key, according to the authors, to the claim that LLMs have only shallow understanding. These systems, in fact, are trained on language and thus acquire only the small part of human knowledge that is linguistic. The remaining part cannot be learned by the machine. As Browning and LeCun argue, “a system trained on language alone will never approximate human intelligence, even if trained from now until the heat death of the universe,” precisely because linguistic knowledge “is just the wrong kind of knowledge for developing awareness or being a person.” It is easy to see that the “everything else” the authors refer to is “tacit knowledge,” as conceived by Polanyi. And, indeed, the authors write that this is a kind of knowledge that human beings acquire “from exploring the world,” from “social customs and rituals,” and “in the form of precise movements passed on from skilled practitioner to apprentice.” The similarities with Polanyi’s reflections are striking. The authors also state that “the deep nonlinguistic understanding is the ground that makes language useful,” reaffirming the tacit foundation of the human ability to use language, as originally argued by Polanyi. Further, they state that the “broader, context-sensitive kind of learning and know-how is the more basic and ancient kind of knowledge, one which underlies the emergence of sentience in embodied critters and makes it possible to survive and flourish.” By experiencing the world, exploring it, experimenting in it, interacting with it, and, in the case of humans, interacting with culture and other people, living beings gain a “deep understanding” that goes beyond language. As the authors argue, again taking up a Polanyan theme,

Language may be a helpful component which extends our understanding of the world, but language doesn't exhaust intelligence, as is evident from many species, such as corvids, octopi and primates. [...] There is no way to approximate this deep understanding solely through language; it's just the wrong kind of thing. (Browning and LeCun)

And this shows "how little can be known from language alone," reiterating a conclusion already reached by Polanyi.

Faced with the choice of whether to regard the originality and individuality of animals and men as the product of "some ingenious automatic machinery" or as "an independent force operating through the body in combination with the existing machinery of the body" (*PK*, 352), Polanyi strongly advocates the latter. What makes us human, which binds us to all expressions of life but distinguishes us from machines, is the presence of this "independent force" that in other places Polanyi calls the "active center" and that is an expression of the individuality, originality, and personality of every human being. In this respect, Polanyi's philosophy is a philosophy for our time because it is a humanist philosophy, which recognizes that humankind has a special place in the world as a product of a teleologically and finalistically oriented process of emergence. The emergence of humans represents "the appearance of centers of thought and responsibility in the visible words" (*PK*, 428) and "the gradual rise of autonomous centres of decision" (*PK*, 425). In a world where automatic decision-making systems make decisions, replacing human will and the responsibility it implies, Polanyi's thought helps us reestablish the idea that no machine can ever take away from humans those capabilities (making decisions, being responsible, understanding the world) that are the product of millions of years of emergent evolution. Contributing to this understanding of these aspects of Polanyi's thinking is just another of the merits of Paksi and Héder's *Guide to Personal Knowledge*.

REFERENCES

- Browning, J., and Y. LeCun. 2022. "AI And The Limits of Language." *NOÉMA*. August 23, 2022. <https://www.noemamag.com/ai-and-the-limits-of-language/>.
- Paksi, D., and M. Héder. 2022. *Guide to Personal Knowledge: The Philosophy of Michael Polanyi: Tacit Knowledge, Emergence and the Fiduciary Program*. Vernon Press.
- Polanyi, M. 1998 (1962). *Personal Knowledge. Towards a Post-Critical Philosophy*. Routledge, e-book edition.