Review

Problems of Connectionism

Marta Vassallo [1,2,*,†], Davide Sattin [3,†], Eugenio Parati [4] and Mario Picozzi [2]

1   Ph.D. Program in Clinical Experimental Medicine, and Medical Humanities, University of Insubria, Via Ravasi 2, 21100 Varese, Italy
2   Center for Clinical Ethics, Department of Biotechnologies and Life Sciences, University of Insubria, Via O. Rossi 9, 21100 Varese, Italy; mario.picozzi@uninsubria.it
3   Health Directorate, Istituti Clinici Scientifici Maugeri IRCCS, Via Camaldoli 64, 20138 Milan, Italy; davide.sattin@icsmaugeri.it
4   Neurorehabilitation Unit, Istituti Clinici Scientifici Maugeri IRCCS, Via Camaldoli 64, 20138 Milan, Italy; eugenio.parati@icsmaugeri.it
*   Correspondence: mvassallo2@uninsubria.it
†   These authors contributed equally to this work and share first authorship.

Abstract: The relationship between philosophy and science has always been complementary. Today, while science moves increasingly fast and philosophy shows some problems in catching up with it, it is not always possible to ignore such relationships, especially in some disciplines such as philosophy of mind, cognitive science, and neuroscience. However, the methodological procedures used to analyze these data are based on principles and assumptions that require a profound dialogue between philosophy and science. Following these ideas, this work aims to raise the problems that a classical connectionist theory can cause and problematize them in a cognitive framework, considering both philosophy and cognitive sciences but also the disciplines that are near to them, such as AI, computer sciences, and linguistics. For this reason, we embarked on an analysis of both the computational and theoretical problems that connectionism currently has. The second aim of this work is to advocate for collaboration between neuroscience and philosophy of mind because the promotion of deeper multidisciplinarity seems necessary in order to solve connectionism's problems. In fact, we believe that the problems that we detected can be solved by a thorough investigation at both a theoretical and an empirical level, and they do not represent an impasse but rather a starting point from which connectionism should learn and be updated while keeping its original and profoundly convincing core.

Keywords: connectionism; connection theory; cognitive rehabilitation; neuropsychology; philosophy of science

1. Introduction

The relationship between philosophy and science has always been complementary. In fact, like it or not, these two disciplines were born from the same big question: *who are we as humans?*

Philosophy and science were, in fact, born simultaneously among the presocratic philosophers who questioned themselves about nature and its meaning; these kinds of discussions led to mathematical and philosophical considerations about the world and its perfection [1]. Until Socrates' time, philosophical inquiry concerned investigations into numerical principles, celestial motions, and inquiries into the origins and dissolutions of all things. This intellectual tradition also extended to the observation of celestial bodies, including the stars, their spatial relationships, and their trajectories, alongside the examination of celestial phenomena [2]. The early philosophers held the conviction that knowledge could be acquired through direct exploration and analysis of natural phenomena [3]. This collaboration between philosophy and science did not stop with Greek philosophers but has always involved both philosophers and scientists, like Schrödinger,

who actively contributed to this discussion. Erwin Schrödinger's "*Nature and the Greeks*" [4] and "*Science and Humanism*" [5] provide insightful reflections on the relationship between nature, scientific inquiry, and humanistic values. In "*Nature and the Greeks*", Schrödinger explored the philosophical underpinnings of ancient Greek thought and its influence on modern scientific understanding. He highlighted the Greeks' profound appreciation for the interconnectedness of the natural world and their pursuit of fundamental principles governing the universe. On the other hand, in "*Science and Humanism*", Schrödinger delved into the broader implications of scientific discoveries and their impact on human society and culture. In this work, he emphasized the importance of integrating scientific knowledge with ethical and moral considerations, advocating for a holistic approach to understanding and addressing humanity's challenges. Overall, Schrödinger's works offer a synthesis of scientific insights with philosophical and humanistic perspectives, inviting readers to contemplate the deeper significance of scientific exploration within the context of human existence. Another scientist who has been involved in exploring the interconnections between science and philosophy has been Wilson [6], who, from the point of view of a biologist, theorized that the very end-line of science and its intrinsic explanation should reside in the humanities and especially philosophy because only such disciplines could provide a satisfying explanation of science itself. In his work, Wilson advocates for the integration of diverse fields of study and encourages researchers to seek common ground in their pursuit of understanding the complexities of the universe. Today, while science, driven by technical evolution, moves increasingly fast and philosophy shows some problems in catching up with it, it is not always possible to ignore such relationships, especially in some disciplines.

One of these is the case of neurosciences and their relationship with philosophy of mind. The study of brain functioning is decidedly complex and is intrinsically linked to the experimenter himself and to his brain, which is used to analyze human brains from a third perspective. It is undeniable that the fascinating appeal of the philosophy of mind has driven several scientists toward different theories of mind, although this relationship does not always work successfully. For example, consider what Lombroso [7] and phrenology, in general, did by creating a sort of pseudoscience built on false beliefs and bringing on a theory that lacks the structure of a scientific theory. So why was phrenology believed to be true? Probably for historical and cultural reasons, i.e., it was convenient to believe so; in fact, this theory was coherent with the beliefs of the time.

Advances in technology have given us the opportunity to analyze big data, and therefore, theories like Lombroso's probably cannot be followed nowadays by scientists. However, the methodological procedures used to analyze these data are based on principles and assumptions that require a profound dialogue between philosophy and science [8,9].

The philosophy of science is hardly something new, and it has a long and fruitful tradition. A fairly recent noteworthy effort in this sense has been made by Heller [3]. In his work, he explores the intersection of philosophy and science by delving into the compatibility of these disciplines and examining how philosophical inquiry can truly enrich scientific practice. In this framework, the role of philosophy is central to clarifying scientific concepts, solving conceptual issues, and fostering critical reflection within scientific communities [10]. Therefore, philosophical engagement with science is essential for addressing foundational questions and advancing scientific knowledge [11]. For this very reason, this inherent interdisciplinarity we are attributing to connectionism is something that could help it solve its problems because such problems are philosophical, computational, and neuroscientific at the same time, and philosophy has always been concerned with all of these disciplines.

Following these ideas, this work aims to raise the problems that a classical connectionist theory [12] can cause and problematize them in a current cognitive framework, considering both philosophy and cognitive sciences but also the disciplines that are near to them, such as AI and computer sciences [13] and linguistics [14]. We will do so by following the map reported in Figure 1.
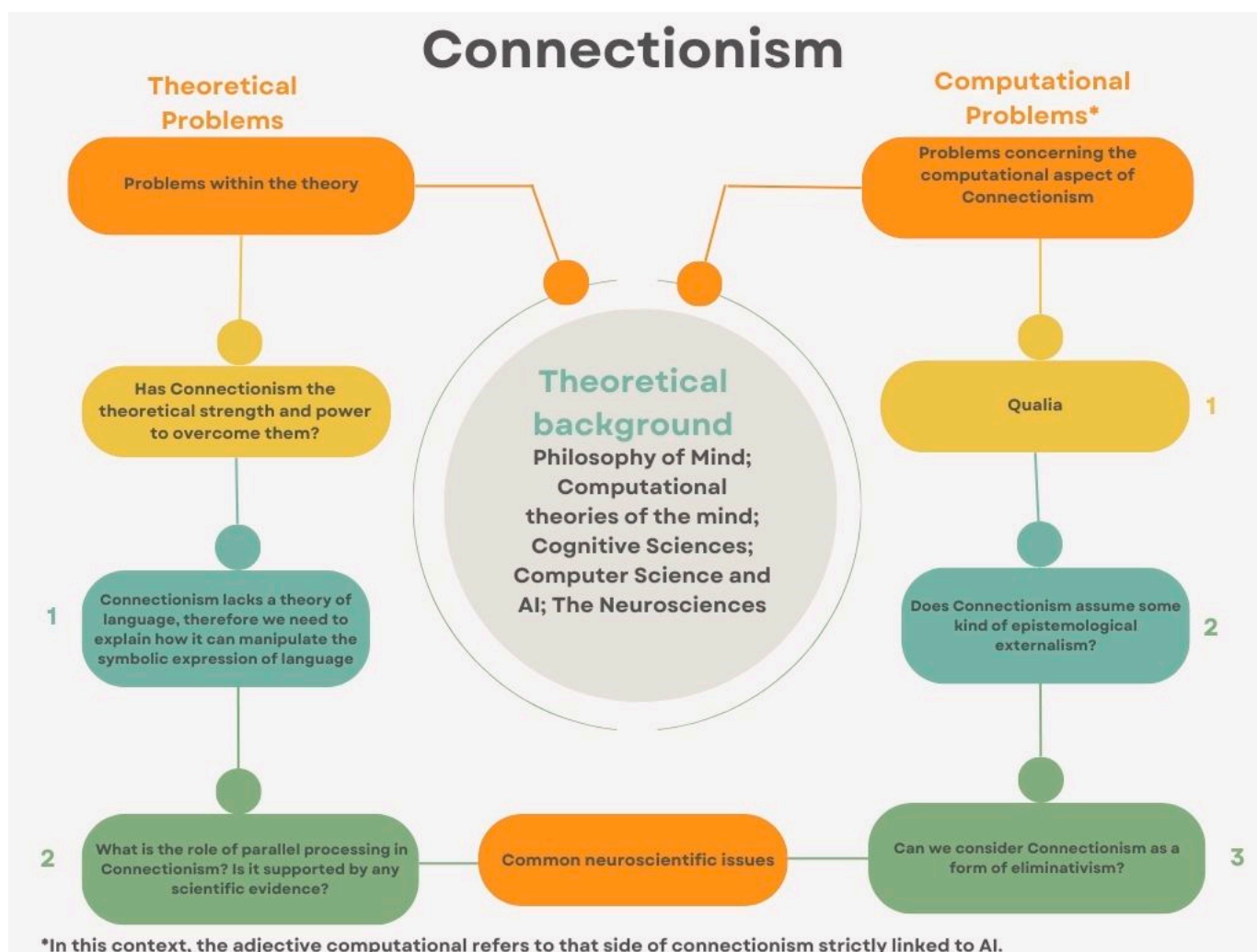
Figure 1. Mind-map of the various ramifications of Connectionisms' problems.

2. Connectionism

Connectionism is a theory that holds in itself some peculiarities from AI, philosophy of mind, and neurosciences and has shown a great deal of effort in this interdisciplinarity process we are advocating here. The basic assumptions of such a theory are the following:

A. Connectionism is a physicalist theory that assumes that our mental states are identical to our brain states [15].

B. Connectionism implies a strong and strict reductionism: all our mental states must be reducible to our brain states, and nothing can be cut off from this reduction [16].

C. In addition, connectionism assumes that artificial neural networks are mirroring systems of our biological neural networks. Therefore, by studying the first, we can uncover a lot about the second [17].

D. This theory is based on AI models and neural models. It uses AI to explain neural functioning [18].

Connectionism is classically considered a form of physicalism or materialism, i.e., the idea that the mind and the brain are identical and every mental state is, in fact, a brain state. While many authors have developed alternative physicalist theories that differ from each other because of very few things, such as views on qualia, e.g., eliminative materialism [19] versus the non-elimination of qualia [20], strictly philosophically speaking, connectionism falls under such a denomination. Such theory differs from other versions of materialism because of its deep attention to AI machine learning systems and their

development and because of its deep link with neuroscience; that is the most remarkable aspect of this theory, as we will see throughout this entire work, along with its profoundly multidisciplinary core and its adherence to new scientific evidence we currently have at hand, that differentiates it deeply from other physicalist theories of mind such as property dualism [21] or behaviorism [22].

Suppose we look at alternative theories to materialism, such as substance dualism [23], i.e., the idea that the mind and the body are two different substances or even functionalism [24], and the classical computational theory of mind [25]. There have been other forms of materialism linked to the neurosciences, and the most famous one is perhaps Searle's biological naturalism. Searle thought that, in the end, the mind would be finally explained by neuroscientific discoveries. However, even if Searle's theory is the closest to connectionism in terms of philosophy of mind, connectionism, in our opinion, has a more complete approach because it has, at the same time, a philosophical backup, a computational one, and a neuroscientific one.

Finally, we are taking into consideration connectionism here, because when other theories of mind lack neuroscientific integration [26] and a consequent philosophical coherence, connectionism has been distinctive. After all, it made a strong effort to implement AI and neuroscience into a philosophical theory.

As technology, AI, and research advance, this theory shows even more evident limits and problems that need to be solved in order to improve it as it is. This means that we do not have to discard the theory in its entirety or pursue some kind of anti-reductionism. But, for the reasons we are going to discuss shortly, it is undeniable that connectionism needs improvement, and it needs to go some steps further, presenting what we define as computational and theoretical problems and some corollary concepts that have emerged in the last few years in the philosophy of mind. In our opinion, these aspects should be analyzed in a connectionist framework, where we consider deep learning to be the younger and more advanced sibling of artificial neural networks. For this reason, we maintain that this theory still has strong theoretical power within itself. For instance, if we consider the aforementioned models as models having some sort of mirroring of the power of human cognition, then we should only consider unsupervised models, as humans—for the most part—learn in an unsupervised setting, i.e., that we are not told what the world is, we just discover it [27].

Therefore, since we have even more sophisticated technologies today that could definitely support an updated connectionist theory, our effort seems even more justified. For instance, some recent research efforts [28] have shown how learning algorithms that try to mimic the way humans learn are not that farfetched and, in fact, could be game-changing in finding a mathematical abstraction of the neurobiological mechanisms that lie at the basis of human brains [29].

3. Computational Problems

Given these premises, it seems necessary to investigate what connectionism has not done, how it can be improved, and its main problems. The following two sections will list some problems that raise computational issues within the theory, that is, problems that create impasses on the computational side of connectionism.

First, the theory has not clearly explained the terms of the comparison between artificial neural networks and biological neural networks. So, a question arises: is this comparison literal? Is it not? If the comparison is literal, this may arise, and actually, it has already arisen [30], raising some issues. These issues concern the fact that if we consider only a connectionist model, this model cannot always make sense. That is, a model without a strong theory keeping it together is somewhat meaningless, at least philosophically speaking. And, unfortunately, connectionism has been focusing on the model (artificial neural nets), leaving aside all that philosophical account that it must take into consideration. Take, for instance, the huge use of neural networks that we have both in

clinical [31] and experimental settings [32]. In most of these cases, the outcome is a strong cluster of evidence, not relying on a strong theory.

Therefore, a big issue concerning connectionism is that the theory stopped considering scientific advancements in both AI and neuroscience. This is a problem because the theory has stopped evolving, probably also due to the lack of philosophers concerned with it, and because today philosophy of language is the main area of investigation in analytic philosophy [33,34].

In parallel, connectionism, even though computational linguistics has made important advancements [35], especially in collaborating with deep learning approaches, has not followed these advancements. When classical connectionist theory is applied to computational linguistics and the philosophy of language, it gives the feeling of being obsolete for two main reasons:

(1) It does not consider the advancements in computational linguistics: since symbol manipulation is the basic trait of computational linguistics, we need to be sure that linguistic symbols are, in fact, reducible to neural connections [16].
(2) The most evident consequence of our former reason is that connectionism does not try to inscribe itself in a theory of language, probably because the symbolic part of the theory remains unclear [36].

This lack of a link between connectionism and a theory of language should be emphasized: if we want to investigate our minds, one special place should be given to our language; that is the utterly important tool we have for conceiving thoughts [37]. As a matter of fact, if in the 1990s a theory of mind could distance itself from a theory of language, this cannot happen today, where the problem of cognition is extended to all areas of cognitive science.

One last major problem that has stopped connectionism from being a coherent theory of mind is linked to its epistemological basis; that is, connectionism inherently declares some sort of externalism, but this has never been stated clearly or theorized properly. However, in this picture, where we are considering the strong reductionist aim of the classical theory [38] we are referring to, it seems clear that internalism cannot be considered a possible epistemological starting point in a theory such as connectionism. But can it be discarded before being discussed? And can some sort of implicit externalism be hypothesized and assumed? That represents an issue. In fact, if externalism is somewhat hinted at by the theory, the explanation concerning it is only constructed referring to artificial neural networks [39] where some explanations, including artificial neural networks (or ANNs) in the light of their relationship with our brains, are needed. In 1996, Paul Churchland [12] suggested one of the main features of connectionism by describing how our brains work with an interesting analogy: our computational ability is described as billions of pixels combined, like a very tall and widescreen composed of several different screens. This is, according to a classical connectionist theory, how the brain can represent the external world. Every part of this system described in this way is somewhat watched by some other parts; this is called parallel processing, and it is what gives us the actual ability to understand what is happening in our brains, at least internally [40]. As it seems obvious, such a connectionist description of our way of capturing and understanding the outside world is not one hundred percent compatible with an externalist epistemological approach, i.e., the knowledge we have of the external world depends uniquely on the external world itself. However, the externalist solution is the one classically hinted at by connectionism. Therefore, there is an evident philosophical gap here that needs to be filled. Another problem concerning parallel processing is that we need at least some neuroscientific evidence guiding us toward the actual acceptance of this instance. Research has shown [41] that for now, it seems plausible but not completely agreed upon and somewhat impossible to demonstrate. Therefore, a question arises: Should we ditch parallel processing? Should we accept it for lack of a better alternative? Should we ditch both parallel processing and externalism? In all fairness, there has been a large use of

parallel processing in research [42–44] However, since the results are partially controversial, the questions above remain open.

It seems, however, that the richness and fullness of humans' cognitive experience cannot be grasped at the current state of the art by current computational [45]. For this very reason, there has been widespread advocacy for neural models that better capture human reasoning by setting precise and definite boundaries on the brain's functions, which such models aim to represent [46].

In conclusion, the epistemological aspect of this theory seems to be a consequence of the theory more than a part of it, and this represents an evident issue that needs to be solved to have a useful, explicative, and complete theory of mind that starts from some connectionist assumptions.

However, it is worth noting that some may argue [47] that since consciousness and human performance are not directly intertwined with AI models, the scientific aim of such models may vary following the explanatory goals of such models, which should not be presupposed as being a mere mirroring of the human mind. It remains to be seen whether the first assumption of the latter statement is true, and it should be one of the goals of a modern connectionist theory to unravel this game-changing point.

4. Theoretical Problems

In this section, we are going to investigate the theoretical problems we can find in connectionism, i.e., problems that create argumentative and philosophical issues inside the theory itself.

The first theoretical problem addressed here is the cognitive puzzle concerning qualia [48]. This issue can be seen as separate from the others we have discussed and are going to discuss because it is a general philosophical puzzle, i.e., more than one theory of mind is concerned with it. The link between the problem of qualia and connectionism is due to the fact that the theory does not state clearly how we should consider qualia. It may seem that total reductionism and eliminative materialism [19] are implicit in a connectionist theory. But, since this solution is problematic for many reasons we are going to discuss shortly, perhaps in order to update connectionism, some other solution may be more efficient and explicative. The classical connectionist theory of mind assumes a reduction in qualia along with everything else [49]. However, the kind of reductionism assumed on qualia by classical connectionism does not quite capture the issue of qualia as a philosophical puzzle for the following reasons, and it may create more problems than the one it solves:

(a)   Reductionism on qualia does not provide an explanation.
(b)   It is still unclear which mental states are designated as qualia and which are not.
(c)   Eliminating something does not solve the problem that arises from the discussion about that something.

Since qualia are strictly bonded to the external world and to our internal perception, an externalist perspective on qualia or a reductionist one could create some problems. Our bodies are the means through which we know, interact with, and understand the outside world we live in. Through our bodies, we are able to know whether a room is too cold, whether we are in pain, and whether we function well or not. Therefore, we could say that *we are* in our body, which is, in turn, in this world. Claiming is, of course, a philosophical challenge both because our personal interpretation of the outside world is obviously subjective; hence, it cannot be experienced or explained by others, but also because it is this experience that allows us to feel and perceive whatever we feel and perceive. In order to understand why this is challenging, we could also add that even if it is true that we do not feel what others feel and we do not, for instance, see what others see, we could reasonably say that since our brains work like others' brains, we must feel, experience, or even see the same things [50]. This answer fits well in a daily, commonsensical kind of inquiry, but if we want to really know whether we experience the same things as other people when we do the same things as them, the answer is not

that straightforward because knowledge and, as a part of it, experience is something often internal to the subject and consequently, inaccessible to others, and it remains uncertain whether the complete elimination or reduction in such experience can solve the problem. On the other hand, we must keep in mind that there have been claims [51] against the idealization of human cognition and the fact that it may not be as ideal as we bound it to be, especially when we try to compare it with computational models. Therefore, an explanation of qualia could even be seen as a *human bug* rather than an ideal functioning model.

In order to solve this issue, we should find a way to test and explain third-person subjectivity, and qualia are the cognitive starting point for doing so.

In conclusion, qualia are still a problem in connectionism, and this problem must be assessed first and then solved as soon as possible.

Another theoretical problem that, considering today's research on AI, seems quite evident is linked to a more ethical and epistemological take on AI. This issue is linked to the effective explainability of AI systems' internal computations and the so-called black-box-like behavior [52].

The problem we are referring to is the so-called explainable AI or XAI problem, and, for the reasons we will see shortly, it cannot be ignored when we draw a connectionist theory of mind. The problem of XAI is, in fact, caused by a black-box-like behavior concerning the internal computation of an AI device, and, of course, deep learning and neural networks are not exempt [53]. This problem arises when the internal computations of a device cannot be retrieved. Hence, a specific outcome is indeed obtained, yet without any clue on how the machine reached such a result or on the internal computations that the machine carried on in order to have such an outcome [54]. By exploring this problem, we also experience yet another gap created by connectionist models, and this gap can be positioned between us, as humans, and the machines we have built and designed. Philosophically speaking, this is a gap in knowledge, i.e., something that creates a partial or complete absence of knowledge. The implications of such a gap are extremely important, so many have advocated for a clearer AI, an explainable AI [55], and more generally, some efforts have been made to try and solve this problem [56]. For instance, this issue also concerns medical diagnostic tools using AI [57] because the need for an explanatory computation is absolutely necessary in a field concerned with human health [58]. For these reasons, many efforts have been made both to shed light on the topic and to propose a new, explainable way to carry on pathological inquiries using XAI [59,60]. Overall, this problem affects not only connectionism as a theory but also a multitude of disciplines and parts of our everyday lives, perhaps even more than we realize [61]. For these reasons, XAI was born and theorized, and it really represents the need, expressed firstly by computer scientists but today also endorsed by many researchers in various fields [62], of having an explanation for the internal computations of an AI system but also of their outcomes. More specifically, this issue is especially relevant within our scope because the black box-like behavior visible in many AI systems today concerns those connectionist networks we are trying to compare to our brains; therefore, if we do not know their internal computations, we may never know our brains.

Philosophy and Neuroscience

Connectionism, in neuroscientific terms, is today known and methodologically applied as connectivity theory [63]. And, when it comes to using the theory as an empirical posit, we have recently seen many studies [64] using such a theory as a taken-for-granted one.

However, let alone the criticalities we have highlighted throughout this entire work and that have been highlighted by many [30,65–68], we also have at hand some neuroscientific evidence that, if interpreted in the present frame, could lead us toward the need for an urgent upgrading of the theory itself, for many important reasons.

First of all, it has been shown that connectivity among children is rather different from connectivity in adults [69], making the connectivity theory something that is not always a fixed piece of theory that one can interchangeably use as a theoretical base for

empiric efforts. For instance, language is understood to be learned by children [70] very differently from adults [71] and it has been postulated that this difference is specifically due to their different connectivity [72], i.e., the different way in which their brains are wired. But language is just a clue of a wider difference between children and adults that seems even more important when it comes to pathological instances such as ADHD [73], DCD [74], neonatal stroke [75], and even low birth weight [76].

This distinction between healthy individuals and pathological ones is also evident in adults, where differences in connectivity have been spotted in cases of insomnia disorder [77], attention deficit disorder, hyperactivity disorder, and bipolar disorder [78], also in Alzheimer's disease [79], and in cases of schizophrenia [80]. For instance, in schizophrenic patients, there has been a reduction in the functional connectivity of the insula and an overall abnormality of the entire area [81], and in Alzheimer's disease and aging, the reduction of both gray and white matters implies a reduction in the brain's connectivity [82].

These are just a few examples taken from the rather large literature on the topic to show that interindividual variety in the brain [83] is way too important to keep the connectivity theory as it is; that is, the theory should not be canceled, but rather it should be innovated and used as a starting point for change. Such change should be devoted to finding, studying, and treating connectivity in each and every brain, i.e., in individual brains, without the extensive generalization that is today the main issue we have spotted in the neuroscientific concept of connectivity [84].

Of course, our critique does not imply that connectivity is tout-court wrong at its core, but the instruments it uses to reaffirm itself may not be that reliable, and for this reason, it could have a different, richer organization than what we are now technically seeing with the available instruments [85]. We hope that by defining the limits of this specific theory, connectionism and, consequently, the theory of neuroscientific connection could improve themselves by considering its limitations, which are probably due to the instruments it uses to obtain the information that is underlying the theory itself; i.e., this is not a problem of methodology but rather an instrument-related problem that we hope will be addressed and solved in the near future.

5. Concluding Remarks

This paper aims to highlight some important problems of connectionism and anticipate some solutions that will be thoroughly investigated in future research. This will be done following the new evolution of both AI and neuroscience.

As we have seen, connectionism is a philosophical theory that has its merits; however, it is also a rather old one, especially because it is not entirely philosophical but concerns equally the neurosciences and computer sciences. For this very reason, and because these disciplines are advancing and moving faster and faster, we need to formulate a new theory based on connectionism that can keep up with such advancements.

In the end, it seems possible that once we solve these theoretical problems, connectionism will be:

(A)   As complete and coherent as possible.
(B)   Ready to move on and outmatch itself.

Point (B) will presumably create a shift in paradigm, giving place to a revolution [49] both in philosophy and neurosciences and giving birth to a kind of neuro-philosophy [86–88] that sees the collaboration between different research as its main strength.

This collaboration and the promotion of deeper interdisciplinarity are necessary for several reasons. For once, neuroscientists are not philosophers, and, of course, philosophers are not always equipped to investigate neuroscientific issues. Also, a two-way collaboration could solve philosophical impasses and ever-lasting problems, and it could provide a strong evidence-based foundation for cognitive philosophy. On the other hand, it could also allow a real evolution in neuroscience, one that has never been seen before. Philosophical support is necessary if we want to have a theory of mind that has consistency and coherency, not

only internally but also externally, i.e., in the relationship between the theory itself and other adjacent disciplines.

Another reason why this collaboration must happen today is linked to AI and its impact in our contemporary times. Deep learning and big data are today widely used to interpret and clarify scientific data. However, they are not exempt from fallacies, and their use raises more issues than we would like to admit. For instance, when the internal computation of a certain model is lost in the model itself, this creates a huge knowledge gap, usually called *black box-like behavior*. This gap cannot be filled only by AI, and it must be solved by the conjoined effort between AI, philosophy, and neuroscience; otherwise, for the reasons we have seen above, it will never be filled. This problem is evident when we consider ontology as the basis of AI: in a system, we must operate following both some kind of epistemology and some kind of ontology. However, an ontology of a model based on an input/output dialogue can only create an epistemic gap, i.e., a gap in knowledge and in the theory of knowledge that is *a priori* assigned to these systems. Therefore, the issue with artificial neural networks is that such systems do not have a precise ontology, i.e., they work very well with coherency, but the human brain is not always a coherent system. Therefore, if we want to understand our brain, we should look at exceptions, and more specifically, at exceptions that are still able to fit into a wholesome system. In fact, if we look for those exceptions, we realize that we have quite a large scientific literature in the neurosciences that considers exceptional cases where some lesions or congenital problems do not correspond to a visible physical impairment [89–91].

However, if we consider AI structures using modeling [92], we realize that their strong ontology-based behavior could be a great starting point to postulate a hybrid model artificial neural network system that could be the computational key to upgrading connectionism; in other words, a strong ontology applied to artificial neural networks could be the game-changing element [47]. Given this theoretical and empirical apparatus, the next philosophical step would be to open up to new perspectives and to re-interpret the connection theory in neuroscience as well as the philosophical structure of connectionism. Finally, we believe that the problems we have analyzed in this work can be solved by a thorough investigation at both a theoretical and an empirical level, and again, they do not represent an impasse but rather a starting point from which connectionism should learn and be updated while keeping its original and profoundly convincing core.

References

1. Amato, F.; López, A.; Peña-Méndez, E.M.; Vaňhara, P.; Hampl, A.; Havel, J. Artificial neural networks in medical diagnosis. *J. Appl. Biomed.* **2013**, *11*, 47–58. [CrossRef]
2. Anderson, W. History and philosophy of science takes form. *Stud. Hist. Philos. Sci.* **2022**, *93*, 175–182. [CrossRef]
3. Holzinger, A.; Goebel, R.; Mengel, M.; Müller, H. (Eds.) *Artificial Intelligence and Machine Learning for Digital Pathology*; Springer International Publishing: Cham, Switzerland, 2020; Volume 12090. [CrossRef]
4. Arredondo, M.M.; Kovelman, I.; Satterfield, T.; Hu, X.; Stojanov, L.; Beltz, A.M. Person-specific connectivity mapping uncovers differences of bilingual language experience on brain bases of attention in children. *Brain Lang.* **2022**, *227*, 105084. [CrossRef]
5. Barttfeld, P.; Petroni, A.; Báez, S.; Urquina, H.; Sigman, M.; Cetkovich, M.; Torralva, T.; Torrente, F.; Lischinsky, A.; Castellanos, X.; et al. Functional connectivity and temporal variability of brain connections in adults with attention deficit/hyperactivity disorder and bipolar disorder. *Neuropsychobiology* **2014**, *69*, 65–75. [CrossRef] [PubMed]
6. William, B. *Mental Mechanisms: Philosophical Perspective on Cognitive Neuroscience*; Routledge: New York, NY, USA, 2008.

7. Brugger, P.; Kollias, S.S.; Müri, R.M.; Crelier, G.; Hepp-Reymond, M.C.; Regard, M. Beyond re-membering: Phantom sensations of congenitally absent limbs. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6167–6172. [CrossRef] [PubMed]

8. Gennaro, C. *Logic in Grammar: Polarity, Free Choice and Intervention*, 1st ed.; Oxford University Press: Oxford, UK, 2013.

9. Chomsky, N. *New Horizons in the Study of Language and Mind*; Cambridge University Press: Cambridge, UK, 2000.

10. Churchland, P.S. Neurophilosophy: The early years and new directions. *Funct. Neurol.* **2007**, *22*, 185–195. Available online: https://pubmed.ncbi.nlm.nih.gov/18182125/ (accessed on 3 March 2024).

11. Churchland Patricia, S. *Neurophilosophy: Toward a Unified Science of Mind-Brain: Toward a Unified Science of the Mind/Brain*; MIT Press: Cambridge, MA, USA, 1986.

12. Churchland, P.S. *Touching a Nerve the Self as Brain*; W.W. Norton&Company: New York, NY, USA, 2013.

13. Churchland, P. *Neurophilosophy at Work*; Cambridge University Press: Cambridge, UK, 2007.

14. Churchland, P.M. *The Engine of Reason, The Seat of the Soul: A Philosophical Journey into the Brain*; MIT Press: Bradford, UK, 1996.

15. Churchland, P.M. *Matter and Consciousness*, 3rd ed.; The MIT Press: Cambridge, MA, USA, 2013.

16. Cowls, J.; King, T.; Taddeo, M.; Floridi, L. Designing AI for Social Good: Seven Essential Factors. *SSRN Electron.* **2019**, *26*, 1771–1796. [CrossRef]

17. Crignon, C.; Lefebvre, D. The long time of dialogue between doctors and philosophers. *Med. Sci. M/S* **2020**, *36*, 1068–1073. [CrossRef]

18. Davis, E.; Marcus, G. Causal generative models are just a start. *Behav. Brain Sci.* **2017**, *40*, e262. [CrossRef]

19. Davis, E.S.; Marcus, G.F. Computational limits don't fully explain human cognitive limitations. *Behav. Brain Sci.* **2020**, *43*, e7. [CrossRef]

20. Davis, S. (Ed.) *Vancouver Studies in Cognitive Science, Connectionism: Theory and Practice: Vol. II*; Oxford University Press: Oxford, UK, 1992.

21. del Pinal, G. The Logicality of Language: A new take on Triviality, "Ungrammaticality", and Logical Form*. *Noûs* **2019**, *53*, 785–818. [CrossRef]

22. delEtoile, J.; Adeli, H. Graph Theory and Brain Connectivity in Alzheimer's Disease. *Neurosci. Rev. J. Bringing Neurobiol. Neurol. Psychiatry* **2017**, *23*, 616–626. [CrossRef] [PubMed]

23. Dennett, D.C. Quining qualia. In *Consciousness in Contemporary Science*; Marcel, A.J., Bisiach, E., Eds.; Oxford University Press: Oxford, UK, 1988; Available online: https://philpapers.org/rec/DENQQ (accessed on 3 March 2024).

24. Dennett, D.C. Quining Qualia. In *Readings in Philosophy and Cognitive Science*; MIT Press: Cambridge, MA, USA, 1993. [CrossRef]

25. Dominey, P.F.; Inui, T.; Hoen, M. Neural network processing of natural language: II. Towards a unified model of corticostriatal function in learning sentence comprehension and non-linguistic sequencing. *Brain Lang.* **2009**, *109*, 80–92. [CrossRef] [PubMed]

26. Ekhlasi, A.; Nasrabadi, A.M.; Mohammadi, M. Analysis of EEG brain connectivity of children with ADHD using graph theory and directional information transfer. *Biomed. Technik. Biomed. Eng.* **2022**, *68*, 133–146. [CrossRef] [PubMed]

27. Esposito, R.; Bortoletto, M.; Miniussi, C. Integrating TMS, EEG, and MRI as an Approach for Studying Brain Connectivity. *Neurosci. Rev. J. Bringing Neurobiol. Neurol. Psychiatry* **2020**, *26*, 471–486. [CrossRef] [PubMed]

28. Feuillet, L.; Dufour, H.; Pelletier, J. Brain of a white-collar worker. *Lancet* **2007**, *370*, 262. [CrossRef] [PubMed]

29. Floridi, L. *Ethics of Artificial Intelligence Principles, Challenges, and Opportunities*; Oxford University Press: Oxford, UK, 2023.

30. Fodor, J.A.; Pylyshyn, Z.W. Connectionism and cognitive architecture: A critical analysis. *Cognition* **1988**, *28*, 3–71. [CrossRef]

31. Friston, K.J. Functional and effective connectivity: A review. *Brain Connect.* **2011**, *1*, 13–36. [CrossRef] [PubMed]

32. Gallo, M.; Krajňanský, V.; Nenutil, R.; Holub, P.; Brázdil, T. Shedding light on the black box of a neural network used to detect prostate cancer in whole slide images by occlusion-based explainability. *New Biotechnol.* **2023**, *78*, 52–67. [CrossRef]

33. Gatti, U.; Verde, A. Cesare Lombroso: Methodological ambiguities and brilliant intuitions. *Int. J. Law Psychiatry* **2012**, *35*, 19–26. [CrossRef]

34. Genon, S.; Eickhoff, S.B.; Kharabian, S. Linking interindividual variability in brain structure to behaviour. *Nat. Rev. Neurosci.* **2022**, *23*, 307–318. [CrossRef]

35. Goksan, S.; Argyri, F.; Clayden, J.D.; Liegeois, F.; Wei, L. Early childhood bilingualism: Effects on brain structure and function. *F1000Research* **2020**, *9*, 370. [CrossRef] [PubMed]

36. Goździalski, K. Normativity and Ontology of Law in Early Greek Philosophy. *Acta Univ. Lodziensis. Folia Iurid.* **2022**, *100*, 107–125. [CrossRef]

37. Guercio, J.M. The importance of a deeper knowledge of the history and theoretical foundations of behaviorism and behavior therapy: Part 2—1960–1985. *Behav. Anal. Res. Pract.* **2020**, *20*, 174–195. [CrossRef]

38. Hansen, H.A.; Li, J.; Saygin, Z.M. Adults vs. neonates: Differentiation of functional connectivity between the basolateral amygdala and occipitotemporal cortex. *PLoS ONE* **2020**, *15*, e0237204. [CrossRef]

39. Hartshorne, J.K.; Tenenbaum, J.B.; Pinker, S. A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition* **2018**, *177*, 263–277. [CrossRef]

40. Haselager, W.F.G.; Van Rappard, J.F.H. Connectionism, Systematicity, and the Frame Problem. *Minds Mach.* **2013**, *23*, 161–179.

41. Heller, M. How is philosophy in science possible? *Philos. Probl. Sci. (Zagadnienia Filoz. W Nauce)* **2019**, *66*, 231–249. Available online: https://zfn.edu.pl/index.php/zfn/article/view/482 (accessed on 3 March 2024).

42. Ingram, K. *IRIE International Review of Information Ethics Vol. 28 (06/2020) Katrina Ingram: AI and Ethics: Shedding Light on the Black Box 2 Author: IRIE International Review of Information Ethics Vol. 28 (06/2020)*. 2020. Available online: https://informationethics.ca/index.php/irie/article/view/380 (accessed on 3 March 2024).

43. Izadi-Najafabadi, S.; Rinat, S.; Zwicker, J.G. Brain functional connectivity in children with developmental coordination disorder following rehabilitation intervention. *Pediatr. Res.* **2022**, *91*, 1459–1468. [CrossRef]

44. Jacquette, D. Kripke and the Mind-Body Problem. *Dialectica* **1987**, *4*, 293–300. Available online: https://www.jstor.org/stable/42970584 (accessed on 3 March 2024). [CrossRef]

45. Jespersen, K.V.; Stevner, A.; Fernandes, H.; Sørensen, S.D.; Van Someren, E.; Kringelbach, M.; Vuust, P. Reduced structural connectivity in Insomnia Disorder. *J. Sleep Res.* **2020**, *29*, e12901. [CrossRef] [PubMed]

46. Jiménez-García, J.; García, M.; Gutiérrez-Tobal, G.C.; Kheirandish-Gozal, L.; Vaquerizo-Villar, F.; Álvarez, D.; del Campo, F.; Gozal, D.; Hornero, R. An explainable deep-learning architecture for pediatric sleep apnea identification from overnight airflow and oximetry signals. *Biomed. Signal Process. Control.* **2024**, *87*, 105490. [CrossRef]

47. Kim, J. *Physicalism, or Something Near Enough*; Princeton University Press: Princeton, NJ, USA, 2007. [CrossRef]

48. Krotov, D.; Hopfield, J.J. Unsupervised learning by competing hidden units. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 7723–7731. [CrossRef] [PubMed]

49. Kuhn, T.S.; Thomas, S.; Hacking, I. *The Structure of Scientific Revolutions*, 2012th ed.; The University of Chicago Press: Chicago, IL, USA, 1962.

50. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]

51. Lewis, D. Elusive Knowledge. *Australas. J. Philos.* **1996**, *74*, 549–567. [CrossRef]

52. Li, K.; Kadohisa, M.; Kusunoki, M.; Duncan, J.; Bundesen, C.; Ditlevsen, S. Distinguishing between parallel and serial processing in visual attention from neurobiological data. *R. Soc. Open Sci.* **2020**, *7*, 191553. [CrossRef]

53. Lowe, E.J. Non-cartesian substance dualism and the problem of mental causation. *Erkenntnis* **2006**, *65*, 5–23. [CrossRef]

54. Manning, C.D.; Clark, K.; Hewitt, J.; Khandelwal, U.; Levy, O. Emergent linguistic structure in artificial neural networks trained by self-supervision. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 30046–30054. [CrossRef]

55. Marcus, G.F. Rethinking Eliminative Connectionism. *Cogn. Psychol.* **1998**, *37*, 243–282. [CrossRef]

56. Marcus, G.F. Connectionism: With or without rules? Response to J.L. McClelland and D.C. Plaut. *Trends Cogn. Sci.* **1999**, *3*, 168–170. [CrossRef]

57. Marcus, G.; Marblestone, A.; Dean, T. The atoms of neural computation. *Science* **2014**, *346*, 551–552. [CrossRef] [PubMed]

58. Mcclelland, J.L. The Place of Modeling in Cognitive Science. *Top. Cogn. Sci.* **2009**, *1*, 11–38. [CrossRef] [PubMed]

59. Montavon, G.; Samek, W.; Müller, K.R. Methods for interpreting and understanding deep neural networks. *Digit. Signal Process. A Rev. J.* **2018**, *73*, 1–15. [CrossRef]

60. Morawetz, C.; Riedel, M.C.; Salo, T.; Berboth, S.; Eickhoff, S.B.; Laird, A.R.; Kohn, N. Multiple large-scale neural networks underlying emotion regulation. *Neurosci. Biobehav. Rev.* **2020**, *116*, 382–395. [CrossRef]

61. Northoff, G. What Is Neurophilosophy? A Methodological Account. *J. Gen. Philos. Sci./Z. Für Allg. Wiss.* **2004**, *35*, 91–127. Available online: https://www.jstor.org/stable/25171275?searchText=neurophilosophy&searchUri=/action/doBasicSearch?Query=neurophilosophy&ab_segments=0/basic_search_gsv2/control&refreqid=fastly-default:4e8711ce1f3f03e0afdceaf704776883 (accessed on 3 March 2024). [CrossRef]

62. Osada, T.; Ogawa, A.; Suda, A.; Nakajima, K.; Tanaka, M.; Oka, S.; Kamagata, K.; Aoki, S.; Oshima, Y.; Tanaka, S.; et al. Parallel cognitive processing streams in human prefrontal cortex: Parsing areal-level brain network for response inhibition. *Cell Rep.* **2021**, *36*, 109732. [CrossRef] [PubMed]

63. Páez, A. The Pragmatic Turn in Explainable Artificial Intelligence (XAI). *Minds Mach.* **2019**, *29*, 441–459. [CrossRef]

64. Parks, E.L.; Madden, D.J. Brain connectivity and visual attention. *Brain Connect.* **2013**, *3*, 317–338. [CrossRef]

65. Pasquale, F. *The Black Box Society: The Secret Algorithms That Control Money and Information*; Harvard University Press: Cambridge, MA, USA, 2015.

66. Pedreschi, D.; Giannotti, F.; Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F. Meaningful Explanations of Black Box AI Decision Systems. *Proc. AAAI Conf. Artif. Intell.* **2019**, *33*, 9780–9784. [CrossRef]

67. Pini, L.; Pievani, M.; Bocchetta, M.; Altomare, D.; Bosco, P.; Cavedo, E.; Galluzzi, S.; Marizzoni, M.; Frisoni, G.B. Brain atrophy in Alzheimer's Disease and aging. *Ageing Res. Rev.* **2016**, *30*, 25–48. [CrossRef]

68. Prabhu, S.P. Ethical challenges of machine learning and deep learning algorithms. *Lancet. Oncol.* **2019**, *20*, 621–622. [CrossRef]

69. Pretzel, P.; Dhollander, T.; Chabrier, S.; Al-Harrach, M.; Hertz-Pannier, L.; Dinomais, M.; Groeschel, S. Structural brain connectivity in children after neonatal stroke: A whole-brain fixel-based analysis. *NeuroImage. Clin.* **2022**, *34*, 103035. [CrossRef] [PubMed]

70. Putnam, H. The Nature of Mental States. In *Readings in Philosophy of Psychology, Volume I*; Block, N., Ed.; Harvard University Press: Cambridge, MA, USA, 1980; pp. 223–231. [CrossRef]

71. Refinetti, R. Philosophy of science and physiology education. *Am. J. Physiol.* **1997**, *272*, S31–S35. [CrossRef] [PubMed]

72. Remme, M.W.H.; Bergmann, U.; Alevi, D.; Schreiber, S.; Sprekeler, H.; Kempter, R. Hebbian plasticity in parallel synaptic pathways: A circuit mechanism for systems memory consolidation. *PLoS Comput. Biol.* **2021**, *17*, e1009681. [CrossRef]

73. Roy, A. Is the connectionist notion of subconcepts flawed? In Proceedings of the International Joint Conference on Neural Networks, Barcelona, Spain, 18–23 July 2010. [CrossRef]

74. Ruddick, W. Can Doctors and Philosophers Work Together? *Hastings Cent. Rep.* **1981**, *11*, 12. [CrossRef] [PubMed]

75. Sato, J.; Safar, K.; Vandewouw, M.M.; Bando, N.; O'Connor, D.L.; Unger, S.L.; Taylor, M.J. Altered functional connectivity during face processing in children born with very low birth weight. *Soc. Cogn. Affect. Neurosci.* **2021**, *16*, 1182–1190. [CrossRef]

76. Schneider, S. LOT, CTM, and the elephant in the room. *Synthese* **2009**, *170*, 235–250. [CrossRef]

77. Erwin, S. Nature and the Greeks. *Sci. Soc.* **1955**, *19*, 186–189.

78. Schröedinger, E. *Science and Humanism: Physics in Our Time*; Cambridge University Press: Cambridge, UK, 1951. Available online: https://books.google.it/books/about/Science_and_Humanism.html?id=P3GdnQEACAAJ&redir_esc=y (accessed on 3 March 2024).

79. Seung, S. *Connectome: How the Brain's Wiring Makes us Who We Are*, 2012th ed.; Penguin Books: London, UK, 2012.

80. Sheffield, J.M.; Rogers, B.P.; Blackford, J.U.; Heckers, S.; Woodward, N.D. Insula functional connectivity in schizophrenia. *Schizophr. Res.* **2020**, *220*, 69–77. [CrossRef]

81. Smythies, J.R. Neurophilosophy. *Psychol. Med.* **1992**, *22*, 547–549. [CrossRef]

82. Surianarayanan, C.; Lawrence, J.J.; Chelliah, P.R.; Prakash, E.; Hewage, C. Convergence of Artificial Intelligence and Neuroscience towards the Diagnosis of Neurological Disorders—A Scoping Review. *Sensors* **2023**, *23*, 3062. [CrossRef]

83. Świercz, P. The Problem of the Status of Harmony in Pythagorean Philosophy. *Horyzonty Polityki* **2019**, *10*, 105–118.

84. Tabor, W. A dynamical systems perspective on the relationship between symbolic and non-symbolic computation. *Cogn. Neurodyn.* **2009**, *3*, 415–427. [CrossRef] [PubMed]

85. Tian, Y.; Zalesky, A.; Bousman, C.; Everall, I.; Pantelis, C. Insula Functional Connectivity in Schizophrenia: Subregions, Gradients, and Symptoms. *Biol. Psychiatry. Cogn. Neurosci. Neuroimaging* **2019**, *4*, 399–408. [CrossRef] [PubMed]

86. Wani, N.A.; Kumar, R.; Bedi, J. DeepXplainer: An interpretable deep learning based approach for lung cancer detection using explainable artificial intelligence. *Comput. Methods Programs Biomed.* **2024**, *243*, 107879. [CrossRef] [PubMed]

87. Watson, J.B. Psychology as the Behaviorist Views It. *Philos. Rev.* **1913**, *20*, 158–177. Available online: https://philpapers.org/rec/WATPAT-21. (accessed on 3 March 2024). [CrossRef]

88. Wiesen, D.; Bonilha, L.; Rorden, C.; Karnath, H.O. Disconnectomics to unravel the network underlying deficits of spatial exploration and attention. *Sci. Rep.* **2022**, *12*, 22315. [CrossRef]

89. Wilson, E.O. *Consilience: The Unity of Knowledge*; Random House: New York, NY, USA, 1998.

90. Yu, F.; Jiang, Q.J.; Sun, X.Y.; Zhang, R.W. A new case of complete primary cerebellar agenesis: Clinical and imaging findings in a living patient. *Brain J. Neurol.* **2015**, *138 Pt 6*, e353. [CrossRef]

91. Zeki, S.; Shigihara, Y. Parallel processing in the brain's visual form system: An fMRI study. *Front. Hum. Neurosci.* **2014**, *8*, 506. [CrossRef]

92. Torres, N.V.; Santos, G. The (mathematical) modeling process in biosciences. *Front. Genet.* **2015**, *6*, 169934. [CrossRef]