



Does whole brain emulation entail the emulation of mental disorders?

Roman KRZANOWSKI*

Jacob KRZANOWSKI**

ABSTRACT

This paper argues that in order to replicate all of the human brain's functions in some physical substrate through a process known as whole brain emulation (WBE), we also need to replicate the human brain's capacity for mental disorders. Our argument is based on the assumptions and technical definition of WBE put forward by Anders Sandberg and Nick Bostrom. The paper takes a narrow, focused perspective in that the problem of WBE discussed here is not one of psychiatry, the philosophy of the mind, or psychology but rather one of technology and the philosophy of technology. Thus, issues associated with psychiatry, philosophy of mind, or psychology are not addressed here. The argument presented in this paper makes three assumptions: (1) All mental functions are functions of the brain. (2) The human brain has a propensity for mental disorder. (3) A whole brain emulation will by definition replicate all functions of the human brain because it exactly models the brain it is based upon. These leads to the conclusion that WBE will have the same propensity for mental disorder as a human brain. This argument was found to be valid and sound, based on the given assumptions. The authors therefore propose that the potential for WBEs to develop mental disorders, currently considered at best a marginal issue in WBE research, should become a mainstream element of the ongoing research into WBE technology.

KEYWORDS

whole brain emulation; mental disorders; AI; Andres Sandberg; Nick Bostrom; philosophy of technology; philosophy of AI

* Ph.D., associated professor, Faculty of Philosophy, The Pontifical University of John Paul II, Kraków, Poland. Email: rmkran@gmail.com.

** M.D., consulting psychiatrist, London, GB.

1. INTRODUCTION

The human brain has an innate capacity for abnormal behavior in the form of mental disorders. In other words, a human brain is capable of showing “a behavioral or mental pattern that causes significant distress or impairment of personal functioning” (Bolton, 2008: 6). (Note that a more detailed discussion of the definition of mental disorders follows below.)

This claim is hard to dispute in the light of the available research. The National Alliance on Mental Illness (NAMI) reports that in 2020, 21% of U.S. adults experienced mental illness, with 5.6% of U.S. adults experiencing serious mental illness, while 16.5% of U.S. youths aged 6–17 experienced a mental health disorder in 2016 (*Mental health by the numbers*, 2022). The UK organization MIND reports that in England, one in four people will experience some kind of mental health problem each year (*Fundamental facts about mental health*, 2016). Numerous other publications have reported similar findings as well (e.g., *About mental health*, 2018; *Mental disorders*, 2019; *Mental health facts and statistics*, 2020; Vigo *et al.*, 2022). Thus, we can justifiably conclude that mental disorders occur quite commonly, or as we said earlier, the human brain has “an innate capacity for abnormal behavior in the form of mental disorders”.

So, what is this paper about? It argues that whole brain emulation (WBE), by its very definition, recreates all the functions of a fully functional human brain in some material substrate. As Anders Sandberg and Nick Bostrom put it, “it produces the phenomenological effects of a mind”, and “it will behave in the same way as the original brain” (Sandberg & Bostrom, 2008:7). As such, it will inevitably also replicate the human brain’s capacity for mental disorder. (Note that we assume that this capacity is inherent to the human brain given the abovementioned prevalence of mental disorders in the general population.)

The potential for WBE to also emulate the capacity for mental disorders may seem obvious, yet it has not been explicated or discussed at length in the WBE studies so far (e.g., Sandberg & Bostrom, 2008; Koene, 2012; Koene, 2013; Shanahan, 2015; Eckersley & Sandberg, 2013; Griaciano, 2019; Hickey, 2019).

There is also another aspect to this argument: A WBE is an artefact, something that has been constructed by humans, so like any artefact, it will have flaws. Some of these flaws may be viewed as mental disorders. In this paper’s argument, we therefore need to recognize that a WBE has at least two sources of mental disorders: Certain intrinsic features of the emulated brain and flaws that are introduced while building the WBE. Thus, even if we realize a theoretically perfect artifact with an impractical ten nines availability,¹ mental disorders may still be present in a WBE. Nevertheless, we will not pursue this line of argument further.

¹ For an overview of high availability in artifacts, see Piedat & Hawkins, 2001.

To be clear, a WBE is not a physical facsimile of the brain but rather a recreation of its functions. It is a functional copy in that “it produces the phenomenological effects of a mind”, and “it will behave in the same way as the original brain”. Thus, when we talk about WBE, we are not talking about copying the physical structure of the brain but rather replicating its functions, specifically those functions that produce “the phenomenological effects of a mind”. The assumptions behind, and the preconditions for, WBE are defined in more detailed later.

Why should we study the potential of a WBE to develop mental disorders? The existing studies of WBE, as pointed out earlier, have failed to explore the potential for mental disorders and instead focused on its realizability (e.g., Sandberg & Bostrom, 2008; Shanahan, 2015; Eckersley & Sandberg, 2013; Griaciano, 2019; Hickey, 2019).² If WBE constructs could start demonstrating disruptive behavior, however, we should at the very least be prepared for it, not least because WBE systems in future may serve as our companions, teachers, managers, political leaders, personal advisors, and partners, possibly even controlling some aspects of our lives (e.g., *ITT 2018–2023 technical annex*, 2018). We therefore believe that the potential for mental disorders in WBEs should be studied as part of the ongoing research into WBE technology.

We start by defining the scope and limitations of the discussion in the following section. In Section 3, we define some WBE concepts based on the works of Sandberg and Bostrom (Sandberg & Bostrom 2008), Patrick D. Hopkins (Hopkins, 2012), Sandberg (Sandberg, 2013), and Murray Shanahan (Shanahan, 2015). We present some assumptions that underlie WBE and describe nine levels (main levels and sublevels) of a WBE’s realization, as presented by Sandberg and Bostrom (Sandberg & Bostrom, 2018). In Section 4, we then lay out the main argument of this paper. Finally, in Section 5, we summarize the conclusions of this study.

2. SCOPE AND LIMITATIONS OF THIS PAPER

WBE is a multifaceted problem, and this paper may be read in many ways. To avoid any possible misinterpretations, we must state the objectives and caveats and limit any definitions implied in the paper.

The objective of this paper is to evaluate an argument that is not presented or discussed in the current publications about WBE (e.g., Sandberg & Bostrom, 2008; Shanahan, 2015; Eckersley & Sandberg, 2013;

² A review of the current WBE research topics can be found at Foresight Institute, Workshop Whole Brain Emulation 2023. Available at <https://foresight.org/whole-brain-emulation-workshop-2023/>.

Griaciano, 2019; Hickey, 2019). This argument asserts that based on the current technical assumptions underlying WBE, WBE systems will have a propensity for mental disorders, as defined by internationally recognized organizations like the WHO and APA.

We propose a series of caveats for this paper, which are presented below.

First caveat: Our focus is on a WBE of a human brain that functions normally, such that it does not exhibit, or has been diagnosed with, a mental disorder as defined above. Thus, we are not concerned with emulating non-human brains, such as those of primates, dogs, cats, insects, or other non-human biological entities.

Second caveat: We assume that when we talk about the brain and the mind, we are functionally talking about the same thing. In other words, the functions of the brain and the mind are one in the same, as are the disorders of the brain and the mind. In addition, terms like mental disorder, mental illness, mental dysfunction, mental health conditions, and brain disorder may be used here interchangeably. To be more specific, we refer exclusively to functions that “will produce subjective mental states of the same kind that would have been produced by the particular brain being emulated” (Sandberg & Bostrom, 2008: 11). This condition excludes functions of the brain that are not responsible for subjective mental states. The tacit assumption here is that brain functions responsible for “subjective mental states” can be separated from “other brain functions”. We do not challenge this claim here.

Third caveat: We do not discuss the nature of mental disorders or whether the definitions or concepts related to mental disorders are correct. We also do not discuss models of the brain or the brain–mind duality. The WBE-related problem discussed here is not one of psychiatry, the philosophy of cognition, the philosophy of the mind, or psychology but rather one of technology and the philosophy of technology. Furthermore, this paper does not discuss mind uploading, artificial general intelligence (AGI),³ or brain modeling, because they are outside the scope of this study, even if some authors may equate them with WBE.

Fourth caveat: We ignore situations where WBE is applied to an already dysfunctional brain (i.e., one with a diagnosed mental disorder), because such a case would not bring any new insights to WBE studies. In other words, a WBE of a dysfunctional brain would inevitably exhibit dysfunction, because it would be a functionally exact replica of a dysfunctional brain. In addition, this essay is not intended as a general treatment of disorders in artificial systems (i.e., artifacts).

³ This kind of intelligence is also referred to as human intelligence, general human-level intelligence, or artificial general intelligence. See the references for different conceptualizations of AGI (e.g., Mitchell, 2019; Fjelland, 2020), the prospect of general purpose, human-level intelligence (e.g., Marcus, 2022), or a general-purpose ability to broadly generalize to fundamentally new areas (e.g., Cassimatis, Bello, & Langley, 2008).

Fifth caveat: We are also not interested in partial WBE (see the discussion of levels of emulation later in this paper), nor are we concerned with AI systems that mimic the human mind or AI studies in general, even those involving malfunctioning AI systems of any kind. We therefore do not consider mental-like disorders in robots with artificial minds or reasoning systems, which have often been portrayed in sci-fi literature, as being relevant to this discussion.

Sixth caveat: WBE is associated by some authors with the concept of uploading the mind (e.g., Rothblatt, 2012; Hopkins, 2012; Russell, 2019; Laakasu *et al.*, 2022; Watanabe, 2022), immortality (e.g., Gelles, 2009; Hayworth, 2010; Shanahan, 2015; Laakasu *et al.*, 2022), and transhumanism (e.g., Bostrom, 2005; Gelles, 2009; Hopkins, 2012; Shanahan, 2015; Lee, 2019; Juengst & Moseley, 2019; Baelo-Allué & Calvo-Pascual, 2021). However, such matters have no bearing on the presented argument, so they are not discussed here.

Finally, we need to define how we understand mental disorders in this paper: A mental disorder is understood as “a syndrome characterized by clinically significant disturbance in an individual’s cognition, emotion regulation, or behavior that reflects a dysfunction in the psychological, biological, or developmental processes underlying mental functioning” (*DSM-5*, 2013).⁴ The authors stress, however, that while this definition does not capture all mental disorders, it does at least set out some minimal requirements for such disorders. We will not argue here about whether such a definition is appropriate or not from the perspective of psychiatry, psychology, or the philosophy of the mind, because such a definition would be elusive and impossible to formulate.⁵ Thus, this discussion lies beyond the scope of this paper. Indeed, even the current methods of diagnosis for mental disorders are judged to be not entirely reliable, so we would be unlikely to zero on in a “correct

⁴ The APA’s definition of mental disorder (*What is mental illness*, 2022) is similar to the definition of The International Classification of Diseases 11th Revision (*ICD-11*, 2018) and that of the World Health Organization (WHO) (WHO, 2022). The International Classification of Diseases 11th Revision says: “Mental, behavioral and neurodevelopmental disorders are syndromes characterized by clinically significant disturbance in an individual’s cognition, emotional regulation, or behavior that reflects a dysfunction in the psychological, biological, or developmental processes that underlie mental and behavioral functioning. These disturbances are usually associated with distress or impairment in personal, family, social, educational, occupational, or other important areas of functioning” (*ICD-11*, 2018: 387).

⁵ “Efforts to define psychological abnormality typically raise as many questions as they answer. Ultimately, a society selects general criteria for defining abnormality and then uses those criteria to judge particular cases” (Comer, 2014: 5). “It may come as a surprise to you that there is still no universal agreement about what is meant by abnormality or disorder. This is not to say we do not have definitions; we do. However, a truly satisfactory definition will probably always remain elusive” (Hooley *et al.*, 2017: 27).

definition” here (e.g., Thyrer, 2015; Allsopp *et al.*, 2019; University of Liverpool, 2019; Thorton, 2022).⁶ Nevertheless, it is not our aim to pass judgment on psychiatric practices but rather select a plausible, operational definition in which to ground our discussion. This is why we use the definition put forward by the APA (*DSM-5*, 2013).⁷

3. WHOLE BRAIN EMULATION (WBE)

What is whole brain emulation (WBE)? WBE recreates a fully functional brain, such that it is functionally indistinguishable from the original mind (e.g., Koene, 2006; Sandberg & Bostrom, 2008; Hayworth, 2010; Eckersley & Sandberg, 2013; Sandberg, 2013; Shanahan, 2015). For Sandberg and Bostrom, the brain and the mind are essentially the same, so an exact emulation of the brain will also be an exact emulation of the mind. They state the following: “WBE is [...] a 1-to-1 model of brain functions”, “it produces the phenomenological effects of a mind”, and “it will behave in the same way as the original brain”. They further posit that a WBE will be “truly conscious the same way as a normal person”, such that the emulated brain “will produce subjective mental states of the same kind that would have been produced by the particular brain being emulated”, adding that WBE “is correctly described as a continuation of the original mind” (Sandberg & Bostrom, 2008:11).

Is WBE technically feasible? The assumption behind WBE research and this paper is that WBE is technically feasible (e.g., Sandberg & Bostrom, 2008; Hopkins, 2012; Shanahan, 2015), although we must also recognize WBE as “a theoretical technology” (Shanahan, 2015). This means that while there is no obvious technological roadblock preventing a WBE from being realized, our current technology is simply not up to it.⁸ However, one could strongly argue against the feasibility of WBE based on the grounds that the mind is the product of the whole body (i.e., the embedded mind) rather than

⁶ For a recent discussion of, and references to, the reliability of Psychiatric Diagnoses, see Gøtzsche, 2023.

⁷ In fact, arguing over the precise definition of mental disorder in the context of this discussion is immaterial so long as we agree (along with the current definitions of mental illness) that the human brain has a propensity for dysfunctional processes, whatever the causes may be, whether they be social, biological, biochemical, or neural in nature. We may refer to such processes as mental disorders, mental dysfunctions, clinically significant disturbances in cognition, emotional regulation failures, psychosocial behavior disabilities, impairments in functioning, mental illnesses, or various other terms. Nevertheless, we leave the precise definition of what mental disorders are to the psychiatrists and psychologists.

⁸ For example, such an emulation would not break, at least to our knowledge, any known physical laws in the way that a *perpetuum mobile* system would.

just the brain,⁹ so emulating brain functions would not also emulate the mind. Nevertheless, this argument is not pursued any further in this paper.

What are the assumptions behind WBE? A WBE realization depends upon satisfying eight technical assumptions established by Sandberg and Bostrom (Sandberg & Bostrom, 2008) and Shanahan (Shanahan, 2015), at least for the current conceptualization put forward by these authors. Satisfying these assumptions should be seen as Sandberg and Bostrom's conditions for realizing a WBE, or they would not have made these assumptions. These assumptions are:

- 1) The first assumption is that of limited physicalism, meaning that most mental functions are based on physical phenomena, although Sandberg and Bostrom do not explain this assumption further.
- 2) The second assumption is multiple realizability, meaning that "the same mental property, state, or event, can be implemented by different physical properties, states, and events" (Sandberg & Bostrom, 2008:15). In other words, a WBE is a complete emulation of brain functions without any assumptions about the physical substratum of the emulation. It does not matter in what physical system a WBE is implemented or even if it is implemented as a purely software construct, such as a softbot.¹⁰
- 3) The third assumption is that the relevant functions of the brain are Turing-computable.
- 4) The fourth assumption is that WBE does not require a total understanding of the brain's functioning but rather just an understanding of its parts and their interactions.
- 5) The fifth assumption is that a WBE only requires a simulation of the brain from a certain level up, meaning that it is unnecessary to emulate levels below this (e.g., the microscopic) to achieve a successful WBE.
- 6) The sixth assumption is that at the emulation level, the functions of the simulated components are understood.
- 7) The seventh assumption is that at the emulated level, the simulated components can be realized in an operational computer (i.e., Turing computable).
- 8) Finally, the eighth assumption is that a WBE must reproduce the brain's functions but not necessarily all other bodily functions, so a WBE does not require a fully embodied brain.¹¹

⁹ "[W]e think with our whole bodies, not just with our minds" (Kelly, 2017). Of course, what Kevin Kelly is voicing here is the concept of an embedded mind (e.g. Clark & Chalmers, 1998; Anderson, 2003; Shapiro, 2010; Wheeler, 2011; Varela, Thompson, & Rosch, 2016; Kiverstein, 2018).

¹⁰ A softbot is a virtual robot platform or autonomous software agent (e.g., Nwana, 1996; Schermer, 2007; *ITT 2018–2023 technical annex*, 2018).

¹¹ The proposed model of the brain is not necessarily without its controversy, so its assumptions may be incorrect, thereby excluding the possibility of a WBE in the way it is described.

WBE therefore assumes that to emulate the brain's functions, it is not necessary to emulate the complete physical structure of the brain at all levels. There is also no need to emulate all of the brain's processes, because the brain is responsible for regulating and controlling bodily functions, and these are omitted from WBE (assumptions 1, 4, 5, and 8). Furthermore, the medium in which an emulation is implemented is of no importance as long as it supports accurately emulating the brain's functions (2). Moreover, brain functions must be Turing computable, or in other words, they must be reducible to symbolic processing through a set of primitive operations (assumptions 3 and 7). (For more on Turing computability, see, for example, Sipser, 1997; Carter, 2005; Primiero, 2020; De Mol, 2021.) Of course, these represent strong assumptions about the brain's functions and the brain itself, and they do not have to be accepted outside of this discussion, but their validity is not being debated here.¹²

What are WBE's levels of realization? The WBE framework, as proposed by Sandberg and Bostrom, comprises nine levels (including main levels and sublevels) of emulation for the human mind: 1a, 1b, 2, 3, 4, 5, 6a, 6b, and 6c (Sandberg & Bostrom, 2018). We ignore level 4 and below because they only partially realize the functions of the mind. Instead, we focus on those levels that wholly replicate the functionality of the brain, namely level 5 and above. We therefore refer to levels 5, 6a, 6b, and 6c as WBE⁵⁺.

Sandberg and Bostrom's levels of emulation are arbitrary with no relation to any conceptualization of the mind outside of the work on WBE. It is also unclear as to whether any realization of WBE would progress through levels of any sort, let alone those proposed by Sandberg and Bostrom. Nevertheless, for the sake of simplicity, we accept these emulation levels.

Detailed specifications for the levels of emulation are provided by Sandberg and Bostrom (Sandberg & Bostrom, 2008), so there is little point in repeating them here. Instead, we list their main characteristics. Thus, Sandberg and Bostrom's levels of WBE⁵⁺ are defined below:

- At level 5 ("Individual brain emulation"), the created artifact (the WBE) has all the characteristics of the specific functional brain that is being emulated.

Indeed, other concepts of the brain and mind that would not satisfy these assumptions have been proposed and discussed (e.g., Flanagan, 1994; Searle, 2004; Robinson, 2007; Nath, 2009; Tononi, 2008; Oizumi, Albantakis, & Tononi, 2014; Tononi & Koch, 2015; McGinn, 2017). As these models do not assume the possibility of WBE, the question of mental disorders in synthetic minds like WBEs may be a mute one from their perspective.

¹² We need to keep in mind that "While we know a great deal about the biochemistry of neurons and synapses and the anatomical structures of the brain, the neural implementation of the cognitive level — learning, knowing, remembering, reasoning, planning, deciding, and so on — is still mostly anyone's guess" (Russell, 2019: 25). This lack of knowledge does not keep us from discussing WBE and its capacity for mental illness, but it does prevent us from regarding WBE as something that is just around the corner.

- At level 6a (“Person emulation”), the created artifact is able to “perform all the tasks required for some normal human job” (Sandberg & Bostrom, 2008: 11).
- At level 6b (“Mind emulation”), the created artifact is “truly conscious in the same way as a normal human being” (Sandberg & Bostrom, 2008: 11), with all the subjective mental states being reproduced.
- At the highest level of 6c (“Personal identity emulation”), “the emulation is correctly described as a continuation of the original mind” (Sandberg & Bostrom, 2008: 11).

Thus, the WBE⁵⁺ levels assume that a fully functional mind with all its subjective experiences and mental functions will be reproduced in a WBE. Indeed, a fully functional mind at a WBE⁵⁺ level will exhibit, by design, all of the functions of the original (biological) mind. This level of emulation is therefore the focus of this paper.

4. A WBE’S CAPACITY FOR MENTAL DISORDERS: THE ARGUMENT

A WBE⁵⁺ is assumed to be a faithful reproduction of a fully functional mind with all its subjective experiences and mental functions. Thus, a fully functional WBE⁵⁺ mind will, by design, inherit all the capacities of the original mind being emulated. Moreover, as a human mind has the capacity for mental disorders, a fully functional WBE⁵⁺ mind will also have the same capacity.¹³ The argument that a WBE reproduces the brain’s capacity for mental disorders goes as follows:

- (1) All mental functions are functions of the brain.
- (2) The human brain (mind) has a propensity for mental disorders.¹⁴
- (3) A whole brain emulation (at WBE⁵⁺ levels) will, by definition, replicate the functions of the human brain, as a WBE is a 1-to-1 model of a brain’s functions, so it “will produce subjective mental states of the same kind that would have been produced by the particular brain being emulated”.

Thus,

- (C) A whole brain emulation (WBE) is an exact functional copy of a human brain, or in other words, it will have all of its functions, such that “it produces the phenomenological effects of a mind”, and “it will behave in the

¹³ Mental disorders of a WBE construct, as discussed earlier, are defined in the same way as those of the human mind, namely a behavioral or mental pattern that causes significant distress or impairment in personal functioning. In the WBE case, “personal functioning” refers to the functioning of the WBE artifact.

¹⁴ By “propensity”, we refer to a natural tendency to behave in a particular way or, in the case of WBE, a natural tendency to exhibit dysfunctional behavior (for definitions of “propensity”, see *Propensity*, 2022; *Propensity*, n.d.).

same way as the original brain". As such, it will also have the same propensity for mental disorders as the original brain. (See the discussion of this propensity in A2 below.)

Assumption (1) is a basic premise of WBE, and it implies several assumptions about the mind and brain. These assumptions imply physical reductionism by stipulating that the functions of the mind are fully accounted for by the brain. Thus, from this viewpoint, the mind is physical, because its mental functions are reducible to the physical functions of the brain. This assumption is spelled out in the works of Sandberg and Bostrom (Sandberg & Bostrom, 2008) and Shanahan (Shanahan, 2015).

Assumption (2) is self-evident, because, as discussed earlier, the human brain has an obvious propensity for mental disorders. Moreover, under assumption (1), there is no difference between mental and brain functions, although not all brain functions are necessarily mental functions (see Assumption [5] above), so mental functions are brain functions and mental disorders are brain disorders. Thus, if you copy the brain's functions, you copy the mind's functions. To phrase this differently, it seems reasonable to accept that if a biological brain may at some point exhibit a mental dysfunction, a WBE of it will be subject to the same possibility, because by definition, a WBE "will behave in the same way as the original brain" (Sandberg & Bostrom, 2008:7). What may be disputed is the precise definition of what a mental disorder is.

Finally, assumption (3) derives from the definition of a WBE (Sandberg & Bostrom, 2008; Shanahan, 2015), so it is accepted.

Thus, accepting assumptions (1, 2, and 3) leads to the conclusion (C), namely that a WBE, as an 1:1 working copy of a particular brain in non-biological substrate replicates all capacities of the human brain, including its propensity for mental disorders. Arguing from (1, 2, and 3) to (C) is logically valid, and as the three premises are based on the accepted definition of WBE and its own assumptions (1–8), the argument is sound.

The soundness of this argument has to be qualified, however. Firstly, WBE is a technology for the potentially distant future, assuming that it may be realizable at all. Indeed, it is quite possible that a WBE may never be realized, in which case the argument would not be sound. Secondly, our conclusion (C) is only valid and sound under the assumptions detailed by Sandberg and Bostrom (Sandberg & Bostrom, 2008) and Shanahan (Shanahan, 2015). Should any one of these assumptions be proven to be false, the soundness of the argument may also come into question. Nevertheless, these are not issues to explore here.

5. CONCLUSIONS

We presented the argument that a WBE of a normally functioning human brain will have a propensity to develop certain mental disorders, at least at the WBE⁵⁺ levels, just as human brains do. We found this argument to be valid and sound based on the assumptions proposed by Sandberg and Bostrom (Sandberg & Bostrom, 2008) and Shanahan (Shanahan, 2015).

Now, what use is there in speculating about a WBE's potential to develop mental dysfunctions? Studying the speculative concept of WBE may provide some insightful opportunities for better understanding the workings of the human mind and the brain–mind relation. Through the development of WBE systems, we may also be able to learn more about our own mental disorders and mental functions in general, although this may seem like a rather optimistic take on WBE.

Finally, as it has been pointed out, the detailed analysis, in the light of the current research in cognition and neuroscience, of the eight assumptions behind Sandberg's WBE proposal may shed the light on the feasibility of the whole project and the whole idea of the emulation of the brain would have to be reassessed.

BIBLIOGRAPHY

- About mental health.* (2018). Mental Health. CDC Reports. Retrieved from: <https://www.cdc.gov/mentalhealth/learn/index.htm> (10.10.2022).
- Allsopp, K., Read, J., Corcoran, R., & Kinderman, P. (2019). Heterogeneity in psychiatric diagnostic classification. *Psychiatry Research*, 279, 15–22. Retrieved from: <https://doi.org/10.1016/j.psychres.2019.07.005> (10.10.2022).
- DSM-5.* (2013). American Psychiatric Association. Diagnostic and statistical manual of mental disorders. (5th Ed.). DOI: <https://doi.org/10.1176/appi.books.9780890425596>.
- Anderson, M.L. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, 149(1), 91–130. DOI: [https://doi.org/10.1016/S0004-3702\(03\)00054-7](https://doi.org/10.1016/S0004-3702(03)00054-7).
- Bolton, D. (2008). *What is mental disorder? An essay in philosophy, science, and values.* Oxford: Oxford University Press.
- Baelo-Allué, S. & Calvo-Pascual, M. 2021. *Transhumanism and posthumanism in twenty-first century narrative.* New York: Routledge.
- Bostrom, N. (2005). A history of transhumanist thought. *Journal of Evolution and Technology*, 14(1), 1–25.
- Carter, T. (2005). *Introduction to theory of computation.* Complex Systems Summer School. Retrieved from: <https://csustan.csustan.edu/~tom/Lecture-Notes/Computation/computation.pdf> (10.10.2022).
- Cassimatis, N.L., Bello, P., & Langley, P. (2008). Ability, breadth, and parsimony in computational models of higher-order cognition. *Cognitive Science*, 32(8), 1304–1322. DOI: 10.1080/03640210802455175.
- Clark, A. & Chalmers, D. (1998). The extended mind. *Analysis*, 58, 7–19.

- Comer, R.J. 2014. *Abnormal psychology*. (8th Ed.). New York: Worth Publishers.
- De Mol, L. 2021. Turing machines. In: *The Stanford encyclopedia of philosophy*. (E.N. Zalta, Ed.). Retrieved from: <<https://plato.stanford.edu/archives/win2021/entries/turing-machine/>> (10.10.2022).
- Fjelland, R. (2020). Why general artificial intelligence will not be realized. *Humanities and Social Sciences Communications*, 7(10). DOI: <https://doi.org/10.1057/s41599-020-0494-4>.
- Flanagan, O. (1994). *Consciousness reconsidered*. Cambridge: A Brantford Book.
- Fundamental facts about mental health*. (2016). Mental Health Foundation. Retrieved from: <https://www.mentalhealth.org.uk/publications/fundamental-facts-about-mental-health-2016> (10.10.2022).
- Gelles, D. (2009). *Immortality 2.0: a silicon valley insider looks at California's transhumanist movement*. Retrieved from: <https://ce399eugenics.wordpress.com/2010/06/19/immortality-2-0-a-silicon-valley-insider-looks-at-californias-transhumanist-movement/>(10.10.2022).
- Giordana, J.Y., Roelandt, J.L., & Porteaux, C. (2010). La santé mentale des personnes âgées: Prévalence et représentations des troubles psychiques. *Encephale*, 36(3 Suppl.), 59–64. DOI: 10.1016/S0013-7006(10)70018-X. PMID: 20813225.
- Götzsche, P. (2023). Psychiatric diagnoses are not reliable (Part two: Mad in America). In: *Critical psychiatry textbook*. Retrieved from: <https://www.madinamerica.com/2023/03/critical-psychiatry-textbook-chapter-5-psychiatric-diagnoses-are-not-reliable-part-two/> (10.10.2022).
- Griaciano, M. (2019). *What happens if your mind lives forever on the internet?*. Retrieved from: <https://www.theguardian.com/technology/2019/oct/20/mind-uploading-brain-live-forever-internet-virtual-reality> (10.10.2022).
- Hayworth, K. (2010). *Killed by bad philosophy: Why brain preservation followed by mind uploading is a cure for death*. Retrieved from: <http://www.brainpreservation.org> (10.10.2022).
- Hickey, A. (2019). *Whole brain emulation: A giant step for neuroscience*. Retrieved from: <https://www.emergingtechbrew.com/stories/2019/08/15/whole-brain-emulation-giant-step-neuroscience> (10.10.2022).
- Hooley, J.M., Butcher, J.M., Nick, M.K., & Mineka, S. (2017). *Abnormal psychology*. (7th Ed.). Boston: Pearson.
- Hopkins, P.D. (2012). Why uploading will not work, or, the ghosts haunting transhumanism. *International Journal of Machine Consciousness*, 4(1), 229–243.
- ICD-11. (2018). *International classification of diseases 11th revision (ICD-11)*. Retrieved from: <https://icd.who.int/browse11/l-m/en#/http://id.who.int/icd/entity/334423054> (14.07.2023).
- ITT 2018–2023 technical annex. (2018). Istituto Italiano di Tecnologia. Genoa CRL. Retrieved from: <https://www.iit.it/documents/20123/189963/Tech+Annex.pdf/43fe579c-c2c0-c715-e292-c4d39920465a?t=1602744742652> (03.04.2021).
- Juengst, E. & Moseley, D. (2019). Human enhancement. In: *The Stanford encyclopedia of philosophy*. (E.N. Zalta, Ed.). Retrieved from: <https://plato.stanford.edu/archives/sum2019/entries/enhancement/> (22.06.2023).
- Kelly, K. (2017). *The myth of a superhuman AI*. Retrieved from: <https://www.wired.com/2017/04/the-myth-of-a-superhuman-ai/> (12.07.2023).
- Kiverstein, J. (2018). Extended cognition (p.19–32). In: A. Newen, S. Gallagher, & L. de Bruin (Ed.). *The Oxford handbook of 4E cognition*. Oxford: Oxford University Press.
- Koene, R.A. (2006). Scope and resolution in neural prosthetics and special concerns for the emulation of a whole brain. *The Journal of Geoethical Nanotechnology*, 1, 21–29. Retrieved from: https://www.terasemjournals.org/GNJJournal/GN0104/koene_01a.html (22.04.2023).
- Koene, R.A. (2012). Fundamentals of whole brain emulation: State, transition and update representations. *International Journal on Machine Consciousness*, 4(1), 5–21.

- Koene, R.A. (2013). Uploading to substrate-independent minds (pp. 146–156). In: M. More & N. Vita-More (Ed.). *The transhumanist reader: Classical and contemporary essays on the science, technology, and philosophy of the human future*. (1st Ed.). New York: John Wiley & Sons, Inc.
- Laakasu, M., Sundvall, L.J., Drosinou, M., Hannikainen, I., Kunnari, A., Francis, K.B., & Palomäki, J. (2022). Would you exchange your soul for immortality? Existential meaning and afterlife beliefs predict mind upload approval. *PsyArXiv*. DOI: 10.31234/osf.io/ge8t4.
- Lee, N. (2019). *The transhumanism handbook*. Cham: SpingerVerlag.
- Marcus, G. (2022). Artificial general intelligence is not as imminent as you might think. *Scientific American*. Retrieved from: <https://www.scientificamerican.com/article/artificial-general-intelligence-is-not-as-imminent-as-you-might-think1/> (10.05.2022).
- McGinn, C. (2017). *Philosophical provocations*. Cambridge: The MIT Press.
- Mental disorders*. (2019). World Health Organization. Retrieved from: <https://www.who.int/news-room/fact-sheets/detail/mental-disorders> (21.07.2022).
- Mental health by the numbers*. (2022). NAMI. National Alliance on Mental Illness. Retrieved from: <https://www.nami.org/mhstats> (10.10.2022).
- Mental health facts and statistics*. (2020). Mind. Retrieved from: <https://www.mind.org.uk/information-support/types-of-mental-health-problems/statistics-and-facts-about-mental-health/how-common-are-mental-health-problems/> (10.10.2022).
- Milkowski, M. (2020). The computational theory of mind. In: *The Internet encyclopedia of philosophy*. Retrieved from: <https://iep.utm.edu/compmind/#:-:text=The%20Computational%20Theory%20of%20Mind,is%20also%20known%20as%20computationalism.&text=CTM%20is%20often%20understood%20as,cognition%20is%20manipulation%20of%20representation> (03.02.2021).
- Mitchell, M. (2019). *Artificial intelligence: A guide for thinking humans*. London: Penguin.
- Nath, R. (2009). *Philosophy of artificial intelligence*. Boca Raton: Universal Publishers.
- Nwana, H.S. (1996). Software agents: An overview. *Knowledge Engineering Review*, 21(3), 205–244.
- Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0. *PLoS Computational Biology*, 10(5). DOI: 10.1371/journal.pcbi.1003588.
- Piedad, F. & Hawkins, M. (2001). *High availability: Design, techniques, and processes*. Hoboken: Prentice Hall.
- Primero, G. (2020). *On the foundations of computing*. Oxford: Oxford University Press.
- Propensity*. (n.d.). In: *The Britannica dictionary*. Retrieved from: <https://www.britannica.com/dictionary/propensity> (10.10.2022).
- Propensity*. (2022). In: *Collins dictionary online*. Retrieved from: <https://www.collinsdictionary.com/dictionary/english/propensity> (03.02.2021).
- Robinson, D. (2007). *Neuroscience & philosophy*. New York: Columbia University Press.
- Rothblatt, R. (2012). The TERASEM Mind uploading experiment. *International Journal of Machine Consciousness*, 4(1), 141–158.
- Russell, S. (2019). *Human compatible. AI and problems of control*. London: Penguin.
- Sandberg, A. (2013). Feasibility of whole brain emulation (vol. 5, pp. 251–264). In: V. Müller (Ed.). *Philosophy and theory of artificial intelligence. Studies in applied philosophy, epistemology and rational ethics*. Berlin–Heidelberg: Springer. DOI: https://doi.org/10.1007/978-3-642-31674-6_19.
- Sandberg, A. & Bostrom, N. (2008). *Whole brain emulation: A roadmap. Technical report #2008/3*. Oxford: Future of Humanity Institute (University of Oxford). Retrieved from: www.fhi.ox.ac.uk/reports/2008-3.pdf (10.10.2022).

- Schermer, B.W. (2007). *Software agents, surveillance, and the right to privacy: A legislative framework for agent-enabled surveillance*. Leiden: Leiden University Press.
- Searle, J. (2004). *The mind. Introduction*. Oxford: Oxford University Press.
- Shanahan, M. (2015). *The technological singularity*. Cambridge: The MIT Press.
- Shapiro, L. (2010). *Embodied cognition*. London: Routledge, Taylor & Francis.
- Sipser, M. (1997). *Introduction to the theory of computation*. Boston: The MIT Press.
- Skoog, I. (2011). Psychiatric disorders in the elderly. *The Canadian Journal of Psychiatry*, 56(7), 387–397. DOI: 10.1177/070674371105600702. PMID: 21835102.
- Thornton, T. (2022). *Mental illness*. Cambridge: Cambridge University Press.
- Tononi, G. & Koch, C. (2015). Consciousness: here, there and everywhere?. *Philosophical Transactions of the Royal Society, B 370*, 20140167. DOI: <http://dx.doi.org/10.1098/rstb.2014.0167>.
- Tononi, G. (2008). Consciousness as integrated information: A provisional. *The Biological Bulletin*, 215, 216–242.
- Thyer, B.A. (2015). The DSM-5 definition of mental disorder: Critique and alternatives (pp. 45–65). In: B. Probst (Ed.). *Critical thinking in clinical assessment and diagnosis. Essential clinical social work series*. Cham: Springer. DOI: https://doi.org/10.1007/978-3-319-17774-8_3.
- University of Liverpool. (2019). Psychiatric diagnosis ‘scientifically meaningless’. *ScienceDaily*, 8 July. Retrieved from: www.sciencedaily.com/releases/2019/07/190708131152.htm (10.07.2022).
- Varela, F., Thompson, E., & Rosch, E. (2016). *The embodied mind: Cognitive science and human experience*. Cambridge: The MIT Press.
- Vigo, D., Jones, L., Atun, R., & Thornicroft, G. (2022). The true global disease burden of mental illness: still elusive. *The Lancet Psychiatry*, 9(2), 98–100.
- Watanabe, M., (2022). Rethinking our consciousness: An approach to a scientifically feasible seamless mind-uploading. *Research Features*, 145. DOI: 10.26904/RF-145-3507959275.
- Wheeler, M. (2011). Embodied cognition and the extended mind (pp. 220–238). In: J. Garvey (Ed.). *The continuum companion to philosophy of mind*. London: Continuum Press.
- What is mental illness*. (2022). APA. Retrieved from: <https://www.psychiatry.org/patients-families/what-is-mental-illness> (10.07.2023).
- WHO. (2022). *Mental disorders*. Retrieved from: <https://www.who.int/news-room/fact-sheets/detail/mental-disorders> (10.09.2022).