Dr hab. Marek Pokropski, prof. UW   Warsaw, 25.04.2025

Faculty of Philosophy,

University of Warsaw

Review of Dr. Hajo Greif's Academic Achievements in the Habilitation Process (The Pontifical University of John Paul II in Krakow)

Dr. Hajo Greif earned (with distinction) his Magister Artium (MA) in philosophy from Johann Wolfgang Goethe University, Frankfurt, in 1997. His MA thesis focused on constructivism in the philosophy of science. He also studied sociology and cultural anthropology as minors. After graduation, he left academia to work in the media sector. Dr. Greif earned a doctorate in philosophy from Technische Universität Darmstadt in 2004 on the basis of his dissertation "Wer spricht im Parlament der Dinge? Über die Idee einer nicht-menschlichen Handlungsfähigkeit" ("Who Speaks in the Parliament of Things? On the Idea of Non-human Agency"). His dissertation, developed during a three-year fellowship at the DFG-funded "Technology and Society" graduate school, was later published as a monograph by Paderborn/Brill, in 2005.

Dr. Greif's academic career spans multiple institutions and research fields. From 2000 to 2004, he was a DFG-funded doctoral fellow at TU Darmstadt's "Technology and Society" program. During 2003/2004, he held a fellowship at IAS-STS in Graz, researching Darwin's evolutionary theory. After part-time teaching, he co-founded the ICT Research Unit in Graz in 2005, gaining experience in project leadership. In 2009, he took a temporary assistant professorship at the University of Klagenfurt, teaching philosophy and Science and Technology Studies (STS). From 2013 to 2015, he was an FWF Erwin Schrödinger Fellow at TU Munich's MCTS, researching smart environments, which formed the basis of awarding him the tile of *Privatdozent* in Austria. He later returned to MCTS as a senior researcher and lecturer, further engaging in STS scholarship. In 2018, he joined the Warsaw University of Technology as a research assistant professor. Balancing roles at MCTS and Warsaw until 2020, his position at Warsaw became full-time and permanent in June 2020.

Also, Dr. Hajo Greif has been actively engaged in academic teaching since 2007, with experience across multiple institutions. At MCTS, he contributed to curriculum development and candidate selection. Over 18 years, he has taught more than 30 courses, including six in English,

covering topics such as philosophy of science and technology, data and society, AI, ethics, and the intersection of philosophy and biology.

After completing his doctorate, he has published a monograph with the renowned publisher Routledge (Taylor and Francis) and several articles published in international peer-reviewed journals. He is the author of 21 book chapters in collective monographs and the author or co-author of 15 articles in peer-reviewed scientific journals. He has also delivered lectures and presentations at international conferences. Additionally, Dr. Greif was principal investigator in two grants: i) 2021–2024 "Turing, Ashby, and 'the Action of the Brain'", Opus19, National Science Centre, ii) 2013–2016 "From Artificial to Ambient Intelligence", FWF Erwin Schrödinger Fellowship.

Dr. Greif is applying for habilitation in Poland despite already obtaining a similar qualification in Austria from Alpen-Adria-Universität Klagenfurt in 2015. Polish authorities (Narodowa Agencja Wymiany Akademickiej) rejected the recognition of the Austrian habilitation. NAWA argued that Austria does not award post-doctoral degrees, making the Austrian habilitation ineligible for equivalence with the Polish habilitated doctor degree. Dr. Greif decided to undergo the Polish habilitation process anew. Dr. Greif's current proposal builds on a revised version of his Austrian "habilitation" thesis published as *Environments of Intelligence. From Natural Information to Artificial Interaction* (Routledge, 2017). This achievement is supplemented by four single-authored and thematically related journal articles. It is important to stress the interdisciplinary character of Dr. Greif's philosophical work, in which he draws on evolutionary biology, theory of information, cognitive sciences, philosophy of mind, and, last but not least, philosophy of technology.

To sum up, in my view, the academic achievements of Dr. Greif are substantial and show that he is an original researcher and author of important monographs and journal contributions. The achievement indicated in the habilitation application, i.e., the monographic book and four articles listed above, is connected to the original work; however, it is distinct enough to represent a new academic achievement rather than a mere resubmission. In the following section, I discuss in detail the publications indicated in the application as the achievement.

1. Evaluation of the achievements indicated in the application

Dr. Greif's habilitation proposal builds on a revised version of his Austrian "habilitation" thesis published as *Environments of Intelligence. From Natural Information to Artificial Interaction* (Routledge, 2017). This achievement is supplemented by four single-authored and related thematically journal articles: i) "What is the Extension of the Extended Mind?" published in

Synthese (2017), ii) "Affording Illusions? Natural Information and the Problem of Misperception" published in *Avant* (2019), iii) "Adaptation and its Analogues: Biological Categories for Biosemantics" published in *Studies in History and Philosophy of Science* (2021), iv) "Likeness-Making and the Evolution of Cognition" published in *Biology & Philosophy* (2022). Whereas the paper in *Synthese* was published before the book, and contains an early version of one of its chapters, the rest of the articles were published after and develop ideas from the monograph.

This collection of interrelated publications, submitted as the habilitation achievement, is presented under the theme "Informational environments as evolutionary scaffolds: the natural history of cognitive artefacts". In these publications, Dr. Greif covers a wide range of topics, including naturalistic and evolutionary accounts of language and mind, extended mind hypothesis, philosophy of information, philosophy of technology (in particular cognitive digital artefacts), philosophical anthropology, and philosophy of cognitive sciences. Dr. Greif states that his main philosophical contributions are the notion of informational environment and "developing an account of artefacts that serve cognitive functions that is embedded in a pluralistic evolutionary narrative and that explores the artefacts' co-evolutionary potential". This is certainly an ambitious and philosophically weighty project that requires interdisciplinary knowledge in philosophy and other disciplines, as well as an ability to provide a clear argumentation when drawing on complex domains and their diverse terminology. I am not sure that Dr. Greif has succeeded in the latter. My general impression is that his monographic book is overloaded with topics which the author tries to integrate with varying results. Also, I am not entirely sure about the scope of the thesis i.e., whether it is specific and concerns digital technology or has the ambition to be a general theory of cognitive artefacts and informational environments. Despite these doubts, the habilitation achievement is certainly an important voice in several ongoing philosophical discussions.

1.1 The monographic book *Environments of Intelligence*; articles: "What is the Extension of the Extended Mind?", "Affording Illusions? Natural Information and the Problem of Misperception".

The monographic book *Environments of Intelligence. From Natural Information to Artificial Interaction* (Routledge, 2017; pp. 218) consists of 2 parts divided into 10 chapters. In this book, Dr. Greif coins the notion of informational environment and discusses how such an environment influences cognition by blurring the boundary between internal mental processes and external cognitive artefacts. The book argues that even relatively simple technologies can reshape human perception and action by altering the way information is accessed and processed. Part I lays the theoretical foundation, discussing the concept of natural information and its role in perception and behavior. In this respect, Dr. Greif examines key theories of natural information (Dretske) and visual perception, including Marr's computationalism and Gibson's ecological psychology.

In chapter 2, Dr. Greif revives and refines the concept of natural information. This chapter prepares the ground for Greif's original conception of informational environments. Dr. Greif discusses the concept of natural information in detail. However, an important concept that, in my view, was discussed too superficially is intentionality. Although Greif takes into account arguments from Dretske, Millikan, Fodor , as well as Dennett's reply to Fodor, and acknowledges that naturalization of intentionality is the right direction, this is a declaration rather than a justified position. In my view, naturalistic accounts of cognition have to address intentionality in order to explain higher cognitive functions such as attention, purposeful action, memory, mental imagery, etc. Furthermore, it would be interesting to see how the notion of affordances, discussed in chapter 3, might help in developing a naturalistic account of intentionality (relation of intentionality and affordances is considered, for example, in Harry Heft (1989) Affordances and the Body: An Intentional Analysis of Gibson's Ecological Approach to Visual Perception, *Journal for the Theory of Social Behaviour*, 19(1), 1–30).

Also, in chapter 3, Dr. Greif discusses how the concept of affordances contributes to the understanding of illusions and misperceptions. This issue is developed further in the article "Affording Illusions? Natural Information and the Problem of Misperception" published in *Avant* (2019). In this article, illusions are understood as phenomena shaped by environmental affordances, rather than simple perceptual errors. Thus, Greif challenges the traditional view that illusions are merely cognitive failures. Instead, the paper argues that illusions can be useful, e.g., helping individuals navigate complex environments by simplifying decision-making. This perspective aligns with contemporary research on heuristics and cognitive biases, reinforcing the idea that illusions are often beneficial rather than purely deceptive. An interesting question worth considering would be how this concept of perceptual illusion explains perceiving and navigating in Virtual Reality (see, e.g., Grabarczyk & Pokropski 2016, "Perception of Affordances and Experience of Presence in Virtual Reality", *Avant* 7(2)). This issue seems especially important since VR technology is discussed in chapter 9. Also, I wonder whether Greif's conception of illusions is not too general, i.e., there are illusions which are not affordance based but, for example, are related to neurological impairments.

In chapter 4, Dr. Greif continues to defend a naturalistic account of information, and in doing so, he connects areas of information theory, biology and philosophy. Accordingly, "natural information is both a nomologically governed relation and domain specific in relevant respects" (p.57). As far as I understand, this nomological relation is a rough regularity rather than a strict law-like relation. The discussed positions and argumentation are interesting but seem one-sided, i.e., there is too little discussion of alternative theories or counterpoints, in particular those that stress the notion of intentionality. For example, some enactivists (e.g., Shaun Gallagher, Dan Hutto) consider intentionality a natural phenomenon related to embodiment and situatedness in an environment. Such a lower-level, enactive intentionality is, according to enactivists, crucial for

the understanding of an organism's functioning in an environment, not necessarily involving the notion of representation, which Greif tries to avoid. Although Dr. Greif identifies himself with 4E (Embodied, Embedded, Enacted, Extended) theories of cognition, he does not discuss intentionality in this context and considers it a higher-level phenomenon that is based on natural information.

The upshot of the monograph is Greif's original concept of informational environment (IE) that relies on the concept of natural information and ecological environment (Gibson). IE differs from the ecological environment, and can be modified by an organism's activities (chapter 5). In chapter 6, Greif defines "the informational environment E of a receiver R is the subset of the proximal events and objects present in his external environment that I-regularly covary with F-conditions at a distal source s and can be detected by R" (p.97). Interestingly, informational environments can contain both natural information (e.g., tracks, weather patterns) and purposefully created signals (e.g., warning calls, territorial markings), that may serve cooperative or deceptive purposes. Last but not least, informational environments can be modified by technology, altering how ecologically relevant information is available, and potentially extending perception and cognition in various ways.

In part II of the book, Dr. Greif elaborates on the issue of technological modifications of IE. He connects his conception of IE framework to 4E cognition theories and introduces an account of cognitive artefacts. In chapter 7, Dr. Greif introduces the extended mind hypothesis (EMH). This chapter relies on an article published earlier in *Synthese* journal in 2017. Interestingly, Greif discusses EMH in the context of evolutionary theories, and utilizing the notion of niche construction, introduces the concept of cognitive niche that is shaped by extensions and cognitive artefacts (Clark, Sterelny). He introduces an original understanding of extensions that are constitutive to an organism in a minimal (constitutive$_W$) and maximal (constitutive$_S$) sense. Extensions which are maximally constitutive are crucial for vital accomplishments of an organism, as well as essential for explaining proper functions of the organism's biological traits.

Moving to cognitive artefacts, Dr. Greif defines them in chapter 8 as artefacts which augment and support cognitive functions in an analogous manner to tools and machines supporting physical activities (p. 130). To put it shortly, they make information more accessible. Furthermore, cognitive artefacts convey information that is either convergent with or isomorphic to natural information. Additionally, cognitive artefacts introduce effects such as new mechanisms of collecting, structuring and presenting information; they may affect the purposes of their users, and may alter their users' informational environments (p. 137). The main example of such a cognitive artefact, that is discussed in chapter 8, is pictures. Drawing on Gibson's ecological approach to perception, Greif argues that pictures do not merely represent reality but extract and preserve certain invariants (stable information patterns) from perception. Other examples of cognitive artefacts discussed in the book are the GPS system or digital computers.

Although in the preliminary chapter 1, Dr. Greif writes, that "if one considers language an artefact, it will be the paradigm of strongly constitutive cognitive artefacts… My main focus in this context, however, will be on a subset of cognitive artefacts whose very existence owes to Turing's work, namely digital computers" (p. 9). Then, in chapter 8, he acknowledges that such narrowness as well as the selection of cognitive artefacts' examples discussed in the book are somehow one-sided – they are mainly examples of digital technologies. This raises the issue about possible generalization of Dr. Greif's theory, e.g., whether defined criteria for cognitive artefacts apply to non-digital technologies. Dr. Greif admits that there are pre-digital technologies that had significant impact on changing human activities, including cognitive ones, such as electric light and mechanical clock, and fulfill the definition of cognitive artefact i.e., that they are constitutive$_W$ and have at least one of the additional effects (p. 147). And then he responds to this skeptical worry arguing that the difference between digital and non-digital technologies lies in "the profoundness and systematicity of the modifications of informational relations to our environments" (p. 148). First, it seems odd to me to think about electric light as a cognitive artefact just because it had essential influence on human activities. Would a candle or laundry machine also count? Second, I found Dr Greif's response weak. There were non-digital technologies that profoundly and systematically modified human informational environment. Think, for example, about writing or analogue technologies such as printing press, a radio, or photography etc.

Dr. Greif acknowledges that cognitive artefacts provide information and discusses conditions for this function, i.e., convergence or isomorphism of the provided information with natural information. It would be interesting to see how these conditions enable the function of representation, or what additional condition is required. Dr. Greif might reply that the notion of representation is explanatorily irrelevant (drawing on, e.g., Chemero, A. (2000). "Anti-representationalism and the dynamical stance". *Philosophy of Science*, 67(4), 625-647). In some cases, it might be true, however, not in every case of cognitive activity (e.g., mental imagery, language, or the broad category of the so called "off-line" cognition). It might also be the case that a concept of representation is tacitly assumed in the book. For instance, when Dr. Greif discusses isomorphism, he might tacitly acknowledge a sort of structural representations, i.e., representational function would be possible because of structural correspondence (similarity) between an artefact and some external object, e.g., a picture and an object in natural environment. Another option would be something like Bickhard's conception of interactive representations (e.g., Bickhard, M. H. (1999). Interaction and Representation. *Theory & Psychology*, 9(4)).

Dr. Greif considers perception related pictures to be an "elementary kind of cognitive artefact" (p. 131) and distinguishes them from images, which are not necessarily perception related. I do not see how this distinction is justified. Furthermore, it is problematic because there are images that are cognitive artefacts. For example, a mathematical graph, considered by Dr.

Greif an image, is certainly a cognitive artefact. However, information that a math graph provides converges (or is isomorphic) to mathematical objects, rather than to natural information. I am curious how Dr. Greif would address this issue and other examples of cognitive artefacts utilized in mathematical cognition.

My general worry in respect to the notion of cognitive artefact is that the definition is wrongly construed, i.e., it is both too broad (e.g., electric light) and too narrow (e.g., it does not include such cases as mathematical graphs and different sorts of symbolic representations that are conventional rather than based on informational convergence or structural similarity). My confusion is even greater because in the article "Adaptation and its analogues: Biological categories for biosemantics" Dr. Greif considers language a cognitive artefact. It might be important to elaborate a clear typology of cognitive artefacts and discuss similarities and differences between different types of them. I suppose that the evolutionary-ecological-4E approach adopted by Dr. Greif is limited in addressing some types.

To sum up, the book examines technologies that challenge the traditional separation between mind and environment. Interestingly, the presented account of cognitive artefacts draws on the notion of natural information and ecological, non-representationalist understanding of cognition. I find this project bold and interesting, however, I am not fully convinced that it was accomplished. It seems to me that the notion of representation (structural, symbolic) would be necessary to address, as well as a need to discuss examples of cognitive artefacts other than digital technologies. Also, I found parts of the book unclear, which might be because of my limited knowledge of evolutionary biology or information theory, but also might result from the style of argumentation. The book would be more accessible to a reader if it contained a clear and extended outline of the argument in the preface. Also, because of the lack of an introduction and a summary in each chapter, it is sometimes unclear what a chapter's objective is and where the argumentation is going. In this respect, Dr. Greif's articles are much better structured and clearly written.

Despite these shortcomings, which certainly do not undermine the whole project, my overall evaluation of the book is positive. Dr. Greif's argumentation is original and contributes to several ongoing debates in philosophy of mind, philosophy of technology, cognitive sciences, to mention just a few. The book covers a variety of topics from different scientific disciplines. This certainly shows that Greif is an interdisciplinary scholar having impressive knowledge.

1.2 Article "Adaptation and its Analogues: Biological Categories for Biosemantics"

In the article "Adaptation and its Analogues: Biological Categories for Biosemantics", Dr. Greif argues that functions of language are linked to cognitive traits but do not serve adaptive, biological

needs. He discusses a hierarchical and a mutualistic reading of co-evolution of linguistic and cultural forms with adaptive cognitive traits, and argues in favor of the latter. This paper contributes to the field of teleosemantics and the naturalistic program in philosophy of mind and language. That being said, despite the fact that this article discusses issues from evolutionary biology and teleosemantics, I do not see its strong connection with Dr. Greif's project about informational environments. Certainly, this paper is an important publication in the renowned journal *Studies in History and Philosophy of Science*, but does not contribute directly to the themes of informational environments and cognitive artefacts.

1.3 Article "Likeness-Making and the Evolution of Cognition"

In the article "Likeness-making and the evolution of cognition" published in *Biology & Philosophy* (2022), Dr. Greif speculates about the origin of cognitive artefacts. Drawing on Ian Hacking's notion of likeness-making, he argues that the development of modern language and thought is rooted in basic abilities of mimesis and imitation. Interestingly, Dr. Greif modifies Hacking's conjectures and claims that "the earliest forms of making and using public likenesses helped to establish stimulus-detached and convention-based modes of reference-making, and therefore testify to an emerging ability of symbol use even before the advent of iconic images." (p. 4-5). Furthermore, Dr. Greif considers this practice essentially embodied, and supports his claims with empirical findings from paleoanthropology. The discussed artefacts (shell and bone engravings of geometric patterns) support the view that selected human cognition traits (such as the mimetic ability) emerged gradually and appeared already in *Homo erectus*.

I am not an expert on these issues but I found the article interesting and thought-provoking. It is an interesting mix of paleoanthropology, co-evolutionary theories of human cognition and 4E theories of cognition, that proves again Dr. Greif's impressive knowledge coming from different research fields as well as his skill to propose original contributions. The article's philosophical import lies in supporting evolutionary thinking about language origin. In particular, in supporting theorizing about origins of language with empirical evidence, its interpretations and explanatory value. Also, Dr. Greif considers whether the Peircean conception of modes of reference should be applied to Lower Paleolithic artefacts, such as the engravings. He proposes a defense of the distinction of modes of reference that relies on the embodied and collective nature of "likeness making" practices. Furthermore, Dr. Greif claims that "the regularities of those patterns, their perceptual recognition and their cultural transmission would prefigure rules and conventions of form, which in turn are necessary ingredients of symbolic reference" (p. 19). I found this hypothesis plausible for iconic reference. The primacy of symbolic reference claim seems more disputable. This is so because of the arbitrariness of symbolic conventions in contrast to iconic conventions which are non-arbitrary. It would be good if Dr. Greif addressed the issue of conventionality of symbols in his argumentation.

2. Conclusion

The high scientific quality of the monograph, which makes a significant contribution to the fields of philosophy of technology, philosophy of mind, naturalism, and cognitive sciences, together with the accompanying articles and the substantial academic achievements in the form of other articles published in reputable journals, chapters in monographs, research grants, and international conferences, collectively testify to Dr. H. Greif's research maturity and scholarly independence. I believe that he has fulfilled the statutory requirements (Article 219 of the Act of 20 July 2018 on Higher Education and Science) for candidates seeking the degree of habilitated doctor, and I recommend that he be admitted to the next stages of the habilitation procedure.