Review in the Habilitation Procedure

of dr. Hans - Joachim Greif

dr hab. Krzysztof Sołoducha, prof. WAT

Basis and Objectives of the Review

In accordance with the applicable academic regulations, the basis for this habilitation procedure is the set of scientific achievements submitted by Dr Hans-Joachim Greif, consisting of five texts listed in the summary of the habilitation thesis and described by the candidate as follows:

"I partly base the application on a published, significantly modified version of my earlier Austrian habilitation thesis, supplementing it with a series of four thematically related journal articles. This set of texts is related, but sufficiently different from the previous thesis that it does not constitute a mere resubmission of the same work after so many years, but documents an independent, new achievement."

The task of the reviewer is to assess whether the academic output of Dr Hajo Greif, as a candidate for the degree of *doktor habilitowany*, fulfills the requirements defined in Article 219, section 1, item 2 of the Law on Higher Education and Science, according to which the habilitation achievement must constitute a significant contribution to the development of the discipline—in this case, philosophy.

Introduction

Dr Hajo Greif is an academic philosopher whose work lies at the intersection of extended and embodied cognition, ecological psychology, biosemiotics, the philosophy of natural information, philosophy of technology, and the history of science. His publications explore the philosophical implications of empirical research in fields such as cognitive science, biology, and the philosophy of technology. Central to his work are questions concerning the nature of cognitive artefacts, the role of information and the environment in shaping cognition, the evolutionary origins of human cognitive capacities through interaction with artefacts, and the conceptual foundations of notions such as affordances, language, and culture.

Based on the submitted materials, Dr Greif's scholarly work—particularly as presented in his monograph Environments of Intelligence: From Natural Information to Artificial Interaction and related publications—constitutes a sophisticated and naturalistically grounded exploration

of cognition, technology, and their entangled relationship with the environment. Drawing on a broad spectrum of cognitive science topics, Greif's approach is marked by an original synthesis of contemporary computational perspectives with insights from classical German philosophy and philosophical anthropology.

Of particular note is Greif's engagement with the long-standing cultural-nature divide, as well as with the classical German distinction between Naturwissenschaften and Kulturwissenschaften, especially as it relates to the neutral notion of the Umwelt and its role in shaping cognition. His work addresses and seeks to transcend the traditional dualism of subject and object—the foundational opposition of classical epistemology. This concern echoes a central theme in German Idealist philosophy, which grappled deeply with the Kantian problem of things-in-themselves (Dinge an sich) and the epistemic limits of the subject in grasping them.

A particularly original conceptual contribution is Greif's development of the idea of "naturalisation"—the process by which cognitive agents reduce environmental ambiguity through situated interaction, thereby reshaping their own frameworks of reference. This concept illustrates the mutual shaping of artefacts and social practices and serves as a guiding thread throughout Greif's philosophical inquiries.

Dr Greif positions his work within the tradition of analytic philosophy, making use of conceptual analysis while distancing himself from essentialist or static metaphysical views of human nature. In contrast to traditional philosophical anthropology, he proposes a contingent, artefact-dependent conception of cognition—one that aligns with philosophical naturalism, understood as the view that all phenomena are subject to natural laws and open to empirical inquiry.

Key Themes and Concepts

Dr Greif conceptualizes cognitive processes as the handling of natural information in continuous interaction with the environment. This perspective is rooted in the traditions of late phenomenology and German philosophical anthropology, particularly in the works of Helmuth Plessner and Jakob von Uexküll. Greif effectively translates these traditions into the terminology of contemporary, naturalistically oriented Anglo-American cognitive science, engaging closely with Ruth Millikan's biosemiotics and James Gibson's theory of affordances.

A central concept in his work is that of cognitive artefacts—defined as systems that process natural information through evolutionarily developed cognitive mechanisms. These artefacts serve to support, extend, or substitute human cognitive capacities, in much the same way that physical tools augment physical abilities. In this theoretical framework, information plays a role in cognitive artefacts that is analogous to the function of directed motion in mechanical devices.

Within the scope of his naturalistic program, Dr Greif explores various perspectives on the relationship between biological and cultural evolution, treating evolution itself as a conceptual bridge between the domains of nature and culture. He particularly favors co-evolutionary and scaffolding models, in which the interactions between biological traits and cultural forms are understood as mutual and non-hierarchical. This perspective supports a comprehensive naturalistic approach to the philosophy of mind and language, highlighting the dynamic and reciprocal relation between organism and artefact.

Given that language, for Greif, is a particularly significant cognitive artefact in terms of environmental interaction with natural information, he draws extensively on Millikan's biosemiotics. He interprets biosemantics as a naturalistic and evolutionary framework for understanding the mind and language, employing the analogy of natural selection to identify structural similarities across biological, cognitive, linguistic, and cultural domains. Greif extends Millikan's approach beyond biological entities to encompass non-biological artefacts and social behaviours, engaging with debates on adaptation, evolutionary pluralism, and selected-for functions in natural and cultural systems.

Greif also offers a thoughtful critique and reconstruction of Gibson's theory of affordances, identifying two principal shortcomings: the underdeveloped notion of "specifying information" and the lack of clarity in distinguishing misperception from perceptual illusion. To address the first issue, he reinterprets Gibsonian information using Fred Dretske's relational theory of information, which defines information as an objective, sender-independent relation, grounded in lawful co-variation between states of the world. This analysis reintroduces the classical epistemological problem of the subject-object relation into contemporary cognitive theory.

Regarding misperception and illusion, Greif insists on a clear conceptual distinction. He argues that perception is not merely the representation of physical properties, but rather an embodied interaction with contextually embedded objects. In some instances, perceptual illusions may serve adaptive functions, helping organisms to correctly perceive affordances within a broader ecological context—even if such perceptions deviate from physical accuracy.

His work is deeply aligned with the framework of 4E cognition—embodied, embedded, enactive, and extended. Echoing Millikan's claim that "reasoning is done in the world, not in one's head," Greif maintains that cognition is essentially relational and environmentally embedded. According to this view, all cognitive processes are dynamically coupled with the environment; cognition is not enclosed within the mind but distributed across organisms and their artefactual surroundings.

In his essay on "likeness-making", Greif investigates the evolutionary role of artefacts in cognitive development, drawing on paleoanthropological evidence to argue that practices of shaping and marking artefacts functioned as material scaffolds for the emergence of symbolic reference and collective meaning-making. He proposes a gradualist, polycentric model of cognitive evolution, rejecting adaptationist explanations in favour of constructivist, interaction-based approaches to the emergence of language and culture. Importantly, he

suggests that imitation—a foundational aspect of tool use—may contribute to the development of theory of mind, rather than presupposing it.

In an effort to ground his philosophical claims in empirical contexts, Greif incorporates interviews with engineers, computer scientists, and ergonomists, and analyses case studies such as evolutionary robotics, social robots, virtual environments (e.g., Second Life), and mixed-reality games (Pokémon Go). These examples illustrate how digital artefacts restructure informational environments and shape human cognitive behaviour.

Greif's ongoing analysis of the relationship between humans and technological artefacts continues a well-established tradition within German philosophy of technology. He investigates how computational artefacts introduce new forms of human-environment interaction, particularly through properties such as interactivity and simulation—features that fundamentally distinguish them from earlier mechanical technologies.

A particularly insightful distinction introduced by Greif concerns the informational roles of cognitive artefacts, especially computational ones, and can be expressed in terms of two complementary modes:

Convergence refers to artefacts—such as augmented reality devices—that extend natural perception by producing signals that align with the structure of natural information. Devices like Geiger counters provide paradigmatic examples, as they make accessible environmental information that lies beyond unaided human perception.

Isomorphism, by contrast, applies to simulations and virtual environments that maintain a structural similarity to the natural systems they represent. Examples include flight simulators or humanoid robots designed to emulate human communicative behaviours. This mode of informational representation emphasizes internal coherence and functional resemblance to real-world conditions.

This convergence/isomorphism distinction transcends the conventional real/virtual dichotomy in human-computer interaction research, offering a more nuanced framework for analysing the epistemic function of computational artefacts.

Additional Research Areas and Organizational Engagement

In addition to his primary research in the philosophy of mind, cognitive science, and philosophy of technology, Dr Hans-Joachim Greif has cultivated several orthogonally related research trajectories that further demonstrate his intellectual breadth and scholarly versatility.

The first of these concerns the **history and philosophy of the cognitive sciences**, with a particular focus on the intersection between computational and evolutionary accounts of cognition. This line of research is most prominently exemplified in his ongoing NCN-funded project, *Turing, Ashby, and "the Action of the Brain"* (ref. 2020/37/B/HS1/01809). Here, Dr Greif examines the respective contributions of Alan Turing and W. Ross Ashby, not as mutually exclusive paradigms, but as co-original perspectives whose differences—particularly

regarding formal versus material modelling—highlight important tensions in the evolution of cognitive science. This work has already resulted in peer-reviewed publications (e.g. Greif 2018; Greif, Kubiak & Stacewicz 2024) and forms the foundation of a forthcoming monograph under contract with Routledge.

A second, complementary research strand addresses the **epistemology of exploratory modelling in artificial intelligence**. Dr Greif analyses how AI has shifted from its original ambition to simulate human cognitive processes towards the development of predictive yet often "epistemically opaque" generative models, particularly in the context of contemporary machine learning. His publications critically examine whether these models should be understood as fundamentally non-explanatory, and to what extent epistemic opacity should be seen as a structural feature rather than a temporary deficiency to be resolved through Explainable AI (XAI) methodologies.

A long-standing interest in the **history of Darwinian evolutionary theory** also forms part of Dr Greif's scholarly portfolio. In earlier work (Greif 2015), he traces the dual intellectual influences on Darwin—namely, Alexander von Humboldt's romantic and holistic *Naturphilosophie* and the mechanistic natural philosophy of Darwin's British contemporaries. Dr Greif argues that Darwin's theory emerged from a productive synthesis of these traditions, with *Naturphilosophie* framing the explanandum and mechanistic approaches supplying explanatory tools. This investigation underpins Dr Greif's endorsement of **evolutionary pluralism** and began during his fellowship at IAS-STS in 2003, culminating in later publications.

Finally, Dr Greif has also contributed to **historical-philosophical scholarship** through a collaboration with a Ludwik Fleck specialist. In their joint work (Jarnicki & Greif 2022), the authors reconstruct Fleck's underacknowledged influence on Thomas Kuhn's *Structure of Scientific Revolutions*. They offer textual and contextual evidence that Kuhn's formative engagement with Fleck during his graduate studies significantly shaped his conceptual development—particularly Kuhn's framing of paradigm shifts and scientific revolutions—despite Kuhn's later reluctance to credit Fleck's impact.

Through these diverse research endeavours, Dr Greif has demonstrated a consistent commitment to both historical and systematic philosophical inquiry. His ability to connect foundational issues in science and technology with broader epistemological and historical narratives significantly enhances his scholarly profile.

Critical Remarks and Open Questions

Despite the breadth and relevance of Dr Greif's analysis, several challenges and open questions remain. The most significant, in my view, concerns the tension between the declared naturalistic, Anglo-Saxon theoretical approach and the distinctively German philosophical style that characterizes his publications. His texts frequently employ dense, conceptually rich language, which may hinder accessibility—particularly for readers who are not native speakers of English or those unfamiliar with the traditions of German philosophical prose.

While the depth of thought is evident, the stylistic register occasionally obscures the argumentative structure and makes engagement time-consuming.

A second observation—offered from the perspective of a scholar who is not a specialist in cognitive science—concerns the connection between the concept of Umwelt (environment) and the neurophysiological specialization of certain brain structures, such as Broca's area and Wernicke's area, which are well known for their roles in language processing. From the standpoint of an environmentally oriented cognitive psychology and evolutionary theory, it remains unclear how such specialized neural mechanisms emerged in tandem with the broader ecological and artefactual dimensions of cognition. A naturalized epistemology, in my view, should aspire to account for such transformations, bridging the gap between abstract consideration of environmental interaction and concrete neurocognitive specialization of brain modules.

Third remark: examples of computational artefacts could be used more explicitly to demonstrate the cognitive consequences of their widespread use, in alignment with the concept of digital dementia. Such cases would lend empirical weight to Greif's broader theoretical claims concerning the mutual shaping of cognitive artefacts and the epistemic capacities of systems that process natural information. In this respect, further integration of case studies or empirical observations like multitasking, FOMO effect, dopamine effect, information overload and so on would be welcome to support the theoretical framework and its real-world applicability. Such an account would substantiate the rejection of synthetic a priori judgments and provide further support for an empirically grounded epistemological framework.

Conclusion

Hajo Greif's works present a compelling and rigorous contribution to naturalistic philosophy of mind, language, and technology. By integrating concepts from biosemiotics, ecological psychology, cultural evolution, and philosophical anthropology, Greif develops a theoretically rich and empirically oriented framework for understanding human cognition in an artefact-saturated environment.

His emphasis on **natural information** and the **co-evolutionary role of artefacts** offers an original and in accordance with current tendencies perspective on the **extended, embedded, and embodied nature** of cognition.

His achievements and publications constitute a interesting contribution to the development of philosophy.

Therefore I support his application for a habilitation degree within Polish academic system.